



**This electronic thesis or dissertation has been
downloaded from Explore Bristol Research,
<http://research-information.bristol.ac.uk>**

Author:

Furze, Jonathan

Title:

The Rational Design of a Functional Solvent-free Liquid Protease

General rights

Access to the thesis is subject to the Creative Commons Attribution - NonCommercial-No Derivatives 4.0 International Public License. A copy of this may be found at <https://creativecommons.org/licenses/by-nc-nd/4.0/legalcode>. This license sets out your rights and the restrictions that apply to your access to the thesis so it is important you read this before proceeding.

Take down policy

Some pages of this thesis may have been removed for copyright restrictions prior to having it been deposited in Explore Bristol Research. However, if you have discovered material within the thesis that you consider to be unlawful e.g. breaches of copyright (either yours or that of a third party) or any other law, including but not limited to those relating to patent, trademark, confidentiality, data protection, obscenity, defamation, libel, then please contact collections-metadata@bristol.ac.uk and include the following information in your message:

- Your contact details
- Bibliographic details for the item, including a URL
- An outline nature of the complaint

Your claim will be investigated and, where appropriate, the item in question will be removed from public view as soon as possible.

The Rational Design of a Functional Solvent-free Liquid Protease



Jonathan Henry Furze

A dissertation submitted to the University of Bristol in accordance with the requirements of the degree of Doctor of Philosophy in the Faculty of Science, School of Chemistry, September 2018.

Word Count: 62,000

Abstract

Recent research into the rational design and synthesis of solvent-free liquid proteins has yielded remarkable results. For example, these novel anhydrous biofluids maintain secondary structure at temperatures in excess of 150 °C, are able to dynamically unfold and refold, and retain enzymatic activity towards small molecules at extreme temperature. Accordingly, the global aim of this thesis was to demonstrate proteolytic activity in the absence of solvent via the synthesis of a solvent-free liquid protease and an analogous protein-based substrate. This involved increasing the positive surface charge density of the enzyme Proteinase K (PK) and the substrate protein *bovine serum albumin* (BSA) through chemical cationization, followed by the electrostatic surface grafting of anionic polymer-surfactant molecules, and subsequent lyophilization and melting. Significantly, this methodology was used to successfully synthesise solvent-free liquids of both PK and BSA for the first time, both exhibiting high concentrations (over 250 mg/ml) and lower than solvation levels of water (fewer than 30 water molecules per protein construct).

Biophysical analysis, including circular dichroism spectroscopy, was undertaken on the modified proteins after each stage of the synthesis, assessing how the surface modifications had altered the conformation and structural thermal stability. High levels of secondary structure were present throughout the synthesis for both proteins, although deviation away from the native-like structure was seen after cationization, which was to some extent recovered after surfactant conjugation and solvent-free liquid formation. This indicated that the polymer-surfactant corona created during the synthesis could replace structurally significant water molecules and replicate the majority of the associated water-protein interactions, resulting in melting points in the vicinity of 40 °C.

Enzymatic kinetic assessments of solvent-free liquid PK and its aqueous precursors were conducted *via* the development of a novel circular dichroism spectroscopy-based enzymatic assay; a spectroscopic technique often utilised to examine the secondary structure of a protein. By tracking the reduction in substrate secondary structure, protease activity towards solvent-free liquid BSA, and its aqueous precursors, was evaluated. In general, cationization of either the PK or BSA was shown to reduce proteolytic activity due to both structural alterations and an increase in electrostatic repulsion between the protease and the substrate. Surfactant conjugation also reduced the proteolytic activity, in this case because of the surfactant molecules increasing steric hinderance of the protease active site. Formation of the solvent-free liquid state again reduced the proteolytic activity, due in part to the large increase in viscosity hampering diffusion of the macromolecules. However, the retention of even low levels of proteolytic activity in the absence of solvent indicated that the dynamical freedom given by the water hydration shell to the protein had also been replicated by the polymer corona. This retention of similar

surface interactions enabled the protease to perform the dynamic conformational fluctuations required for proteolytic activity.

In light of the persistent structure, activity, and extremely high protein concentrations in the solvent-free liquids (ca. 250 mg mL⁻¹), a novel protocell encapsulation platform was developed. Here, the successful encapsulation of highly concentrated protein-polymer conjugate within silica colloidosomes was achieved through the use of solvent-free liquid myoglobin. The resulting hybrid protein-containing colloidosomes were structurally stable, shown to maintain spherical geometry under the vacuum of scanning electron microscopy, and did not exhibit signs of protein construct leakage.

The outcomes of the work presented in this thesis challenge the perceived necessity of water in roles such as stabilization of protein structure and proteolytic function, and also expands the gallery of macromolecular liquids. Moreover, the ability to encapsulate extremely high concentrations of structurally stable protein nano-conjugates paves the way for a whole new class of protocells with complex internal chemistry.

Acknowledgments

Firstly, I would like to thank Professor Stephen Mann FRS to whom I am forever indebted to for giving me the opportunity to undertake a PhD within his lab and under his guidance. Without your academic kicks through inciteful questioning, I would not have accomplished all that I present here within this thesis. To Dr. Adam Perriman who co-supervised me through these four years of learning, for the invaluable guidance and many a discussion over a pint or two, I extend a huge thanks. You permitted me freedom within my work that not only let me venture around the research world with helpful direction but permitted me space to grow and learn about myself along the way. I would also like to thank GSK for funding my Ph.D project and Dr. Christabel Fowler for being my industrial supervisor.

Many thanks need to be handed out to all of those who've helped me over the years, either with experimentation, use of scientific instrumentation, educational discussions, editing assistance, or general academic aid; my thanks go out to Dr. Ross Anderson, Dr. Jean-Charles Eloi, the Wolfson Bioimaging Facility, Dr. Pierangelo Gobbo, Dr. Alex Brogan, Dr. Thomas Farrugia, Dr. Sam Briggs, Dr. Maddy Nichols, Dr. Alex Croot, Dr. Nicolette Moreau, Dr. Matt Mulvee, Ms. Mary Jenkinson-Finch, and Ms. Joanna Sparks. A special thanks needs to go out to Dr. James Fothergill who at the beginning of my PhD gave me this piece of advice that I've fervently adhered to, "work 9-5 and don't work too many weekends, you need a life outside of the lab".

The PhD life isn't for everyone, not to say it isn't amazing, but you need to be a special kind of crazy to want to go through what we go through, and none of it would be possible if it wasn't for the friendships that you bring in with you and for the ones that you make along the way. Of course, I'm forever in the debt to the people of the Centre for Organised Matter Chemistry. Since the start of the five years that I've been in the lab they've been ever welcoming and accepting; they either knew about or were living the PhD struggle so could help in the most accurate of ways. For the many crossword lunches, trips to the pub (and further afield), laughs, jokes, tears, and just general kindness and compassion, I thank you all. Special shout out to my lab girlfriend Dr. Laura Rodríguez-Arco! Also, to Tim Harrison for affording me the opportunity to take part in copious amounts of outreach over the years.

To my bouldering buddies, thank you for the countless hours of training and dangerous adventures, may there be many more as we ascend to greatness. To my Magic mates, towards ever more mana filled hours spent around dining room tables we travel, with drinks and chats for company. To every dancing diva on the scene, thanks for keeping me sane through the long sweaty dance freestyles. To those involved in Shambala shenanigans, you rock and don't ever change. To my chemphys cohort chums, we've been through it all together, and of course – Chemical physics for life, or until death, whichever comes first! To my home homies, you guys have been around for years and I hope we'll always be there for each other,

may the bubbles always bring us home. To Cate, thank you for all your support throughout, I couldn't have done any of this without you. To my friend and ever humorous housemate Eddie, for the years of laughter had, and for all the years of laughter to come, I thank you. To Maddy, you are an inspiration to me in all areas of my life, your determination to complete a task you've set yourself is boundless and it baffles me, for constantly challenging me even if you weren't aware of it, thank you. To Alex, you simply make me want to be a better person so thanks for that! And, to YouTube, your stimulating content has eased the pain of the last four years.

Sam, my brother from another mother, my comrade, my lab and life partner, but most of all my friend. We started together all those years back and have since trotted around numerous cities, played plenty of board games, consumed lakes of whiskey, created and shared all the stories, and watched an ungodly amount of TV shoulder to shoulder. I can't thank you enough for the part you've played so far in my life, and remember – Together, we are fist!

To my family; my mum, dad, and sister. You have all been there since the beginning and have each in your own ways brought me to where I am today and helped shape who I have become. Mum, you've taught me your compassion and patience, and I've done as best as I can to follow in your kind natured footsteps. Dad, you've instilled determination into my life and educated me in the scientific ways since before I can remember. Lydia, you're a constant reminder for me to have fun, to be occasionally frivolous, to enjoy the little things in life, to be a weird little human, and to never forget that music is the key to happiness. You guys are my rocks to which I anchor this weird ship of mine, may we always remain tethered.

For these people and the others that have cared about me over the last few years of toil, you all know how much you mean to me and how much this weighty tome has tested me. I thank you all.

Authors declaration

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's Regulations and Code of Practice for Research Degree Programmes and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

Signed: _____

Date: _____

Contents

Abstract	i
Acknowledgments	iii
Authors declaration	v
Contents	vi
Chapter 1 Introduction	10
1.1 Introduction	11
1.2 Protein structure and function	13
1.2.1 Primary structure	13
1.2.2 Secondary structure	14
1.2.3 Tertiary structure	18
1.2.4 Quaternary structure	20
1.2.5 Stabilization of protein structures	20
1.3 The effect of modification and bio-conjugation on the structure and function of proteins	23
1.3.1 Post-translational modifications	24
1.3.2 Peptide cleavage	25
1.3.3 Bio-conjugation	26
1.4 Solvent-free liquid proteins	32
1.4.1 Synthesis of protein melts	33
1.4.2 Physical characteristics of solvent-free liquid proteins	37
1.4.3 Biological function within solvent-free liquid proteins	43
1.4.4 Solvent-free liquid proteins solvated in ionic liquids	45
1.5 Introduction to proteases	47
1.5.1 General protease mechanisms	47
1.5.2 The protease active site	49
1.5.3 Specific protease mechanisms	50
1.5.4 Proteinase K	54

1.6	References	56
Chapter 2	Materials and methods.....	65
2.1	Introduction	66
2.2	Techniques and associated theory.....	66
2.2.1	Circular dichroism spectroscopy.....	66
2.2.2	Differential scanning calorimetry	72
2.2.3	Michalious-Mentan kinetics	74
2.2.4	Calculation of Gibbs free energy	76
2.3	Materials and methods.....	77
2.3.1	Materials.....	77
2.3.2	Methods	80
2.4	References	88
Chapter 3	Characterisation of solvent-free liquid proteins and their aqueous precursors	89
3.1	Introduction	90
3.2	Results.....	91
3.2.1	Characterisation of cationized proteins	91
3.2.2	Characterisation of surfactant nanoconjugates conjugates.....	115
3.2.3	Characterisation of solvent-free liquid proteins.....	129
3.3	Overall conclusions and future work	152
3.3.1	Conclusions.....	152
3.3.2	Future work	154
3.4	References	155
Chapter 4	Protease activity in the solvent-free liquid state.....	157
4.1	Introduction	158
4.1.1	Protease activity assays	158
4.1.2	Common protease activity assays	160
4.2	The development of a novel circular dichroism spectroscopy-based protease assay	162
4.2.1	CD assay introduction.....	162
4.2.2	Assay functionality.....	162
4.2.3	Assay development: Controls and preliminary results	163

4.3	Results for the aqueous precursors	168
4.3.1	CD assay discussion	168
4.4	Solvent-free liquid protease activity	187
4.4.1	Introduction	187
4.4.2	Solvent-free liquid protein activity results and discussion	188
4.5	Conclusions and future work	193
4.6	References	196
Chapter 5	Encapsulation of solvent-free liquid proteins within colloidosomes.....	198
5.1	Colloidosome introduction	199
5.1.1	Pickering emulsion overview	199
5.1.2	Colloidosomes	201
5.1.3	Colloidosomes as an encapsulation technique.....	203
5.2	Results: Characterisation of solvent-free liquid myoglobin and aqueous precursors	206
5.2.1	Mass spectrometry	206
5.2.2	Charge density measurements using zeta potential	207
5.2.3	Particle size and distribution through dynamic light scattering (DLS)	207
5.2.4	Differential scanning calorimetry	209
5.2.5	Thermogravimetric analysis.....	211
5.2.6	Secondary structure analysis of solvent-free liquid myoglobin and its aqueous precursors 213	
5.2.7	UV/Vis spectroscopy of myoglobin.....	215
5.2.8	Conclusion of characterisation	217
5.3	Results: Colloidosome formation.....	218
5.3.1	Size distribution	218
5.3.2	Location of silica and encapsulated protein	220
5.3.3	Encapsulation of myoglobin nano-conjugate	224
5.3.4	UV-Vis and DR UV-Vis	225
5.3.5	Scanning electron microscopy (SEM) of colloidosomes	227
5.4	Conclusions and future work	228
5.5	References	229

Chapter 6	Conclusions.....	231
6.1	Conclusions	232

Chapter 1 Introduction

1.1 Introduction

The surface modification of proteins, *e.g.* fluorescent tagging¹ and crosslinking², has been shown previously to alter the structure and function of a protein, often in a detrimental manner. However, the modification of proteins is a useful tool in scientific research, allowing proteins to exist, as well as function, in new environments and for new phases to form. One such advance in the field of protein modification is the formation of solvent-free liquid proteins³⁻⁶; a liquid phase where a protein or surfactant construct is in a solvent-free yet active state. The synthesis of this state involves a number of protein surface modifications and a lyophilisation step, each having various structural and functional impacts on the protein. This thesis describes the study and characterisation of two novel solvent-free liquid proteins, solvent-free liquid *bovine serum albumin* and solvent-free liquid proteinase K, investigating the structure retention and high thermal resistance, as well as displaying solvent-free liquid protease activity for the first time.

This first introductory chapter of the thesis involves an outline of the four levels of protein structure (primary, secondary, tertiary, and quaternary), the stabilisation forces involved in the retention of this structure and the overall conformation of a protein. The discussion then turns to look at how the structure of a protein can be altered through natural and synthetic modifications, leading on to how the structure is affected by the changing of solvent (*i.e.* from water to non-aqueous), culminating in an in-depth review of how the removal of all solvent and the formation of a solvent-free liquid protein impacts the structure and function of a protein or enzyme. One of the proteins examined within this thesis is a protease, so this chapter concludes with a short review of proteases covering the role of the catalytic triad/diad, oxyanion hole, binding cleft and a synthetic-mechanistic look at the function of a few general protease groups.

The information provided in chapter 2 is split in to two major sections; the first section is formed of a detailed examination of the theory behind a number of the major analytical techniques used throughout this thesis, and the second section provides the experimental details for the formation of a solvent-free liquid protein and for all other experiments carried out throughout the thesis, such as instrumental set-up and specific experimental parameters.

The work presented in chapter 3 demonstrates how each stage of the solvent-free liquid protein synthesis affects the structure of the investigated proteins, with the chapter being split into three sections relating to the three stages involved in the formation of a solvent-free liquid protein. Firstly, the chemical surface cationization that modifies negative surface point-charges to being positively

charged, secondly the electrostatic addition of anionic-surfactant molecules and finally the lyophilization of the modified protein to remove all solvent water and yield the solvent-free liquid protein product. This work initially focused on the aqueous precursor stages of the process following the modification of two specific proteins with investigations into the particle size, secondary structure and structural thermal behaviour. Similar investigations were then performed on the solvent-free stage of each respective protein, again studying the secondary structure, and structural thermal behaviour, but expanding to the characterisation of some material properties of the liquid.

Following on from the previous chapter, chapter 4 examines the activity of a protease towards a protease protein-substrate; investigating how each stage of the solvent-free liquid protein synthesis on both the protease and the protein-substrate altered the activity of protease towards the substrate. This investigation involved the development of a new protease kinetic assay that tracked the degradation of the substrate secondary structure over the reaction period. Each aqueous precursor was examined, forming a set of nine unique protease – substrate pairings, with a tenth pairing comprised of two solvent-free liquid proteins.

Chapter 5 investigated a potential application of the solvent-free liquid protein systems as model blood cells. Blood cells are known to contain high levels of haemoglobin concentration, an attribute that has yet to be successfully tackled in the field of synthetic red blood cell formation. Through utilising the high protein concentration within solvent-free liquid proteins (over 250 mg/ml) high levels of protein were incorporate within colloidosomes, colloidally stabilised oil/water capsule systems, mimicking this one aspect of a blood cell (high haemoglobin concentration). Characterisation of the protein-containing colloidosomes was conducted, investigating their size distribution, and thermal behaviours and a study conducted to probe the specific location of the solvent-free liquid protein within the capsule.

In chapter 6 conclusions to the work conducted throughout chapters 3-5 are presented, along with potential areas of future work.

1.2 Protein structure and function

Proteins are a large class of biomolecule that are essential for the correct and sustained function of all cells, tissues, and organs within all living animals. There are four distinct levels of structure within most proteins (primary, secondary, tertiary, and quaternary) with an alteration to any stage affecting the structure and function of the protein.

1.2.1 Primary structure

The primary structure of a protein is the linear arrangement of amino acids; a small molecule termed due to one end having an amino group and the other having a carboxylic acid group. Each amino acid has a similar 'backbone' structure (N-C-C) with the central carbon (C_α) having a covalently bound side group (R) that varies for each of the 20 amino acids; the R group giving the name to each amino acid. The C_α is chiral for 19 of the 20 amino acids; the exception being glycine where the R = H making the molecule achiral and as such can exist in two enantiomeric states known as dextrorotatory (D) and levorotatory (L) (Figure 1.1). All amino acids that are synthesised by condons within life on Earth are in the L-isomer state, and as such only L-isomers are found within proteins. At a neutral pH amino acids present as dipolar ions (zwitterions), with the amino group being in a protonated state ($-NH_3^+$) the carboxylic group being in a deprotonated state ($-COO^-$)⁷.

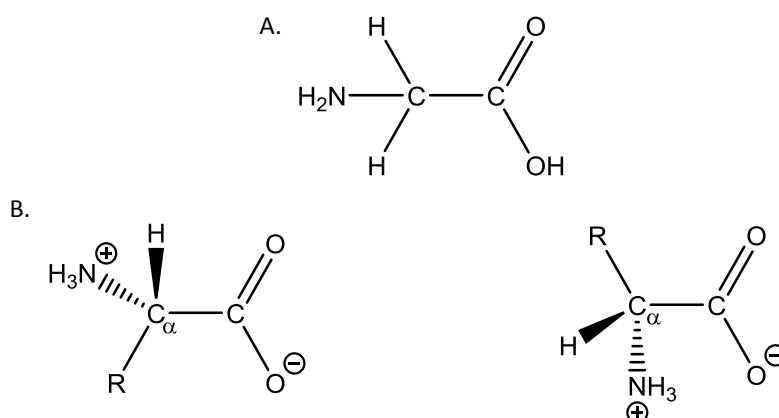


Figure 1.1: **A.** Structure of glycine and **B.** generic structure for an amino acid showing the chirality around C_α .

All amino acids are difunctional and, as such, can covalently couple through a condensation reaction and form a peptide bond which is an amide bond between nitrogen on one amino acid and the carboxylic acid group on the other (Figure 1.2). The peptide bonds are planar in nature due to the resonance structures imparting partial-double-bond character with electrons from the NH lone pair and the π electrons around the C=O double bond. Being planar in geometry the peptide bond can take a *cis* or a *trans* configuration, with the majority taking the *trans* configuration due to the steric hindrance that having both C_α on the same side bond (Figure 1.2). The addition of amino acids to the chain can sequentially be repeated forming a polypeptide chain which starts with an amino terminus (N-

terminus) and ends with a carboxylic terminus (C-terminus). The polypeptide chain forms the primary structure of a protein, with the repeated N-C $_{\alpha}$ -C atoms of each incorporated amino acid (residue) forming the protein backbone.

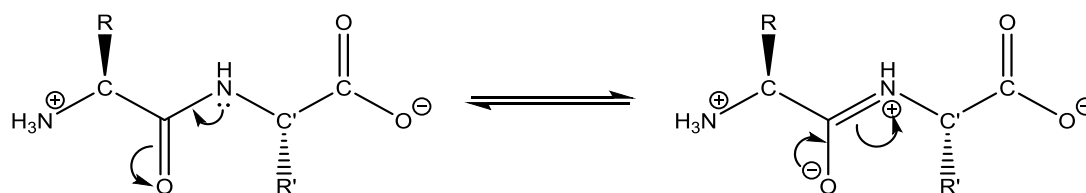


Figure 1.2: Resonance structure for the peptide bond between two generic amino acids.

1.2.2 Secondary structure

In contrast to the rigid partial double-bond nature of the peptide bond (red in Figure 1.3) the covalent bonds that link the C $_{\alpha}$ to the amino and carbonyl group are single bonds⁸. This permits freedom of rotation about these two bonds giving rise to a set of dihedral angles to each C $_{\alpha}$; one between the C $_{\alpha}$ and the amino nitrogen (ϕ) and one between the C $_{\alpha}$ and the carbonyl carbon (ψ) Figure 1.3). Due to the steric repulsion from atoms on the protein backbone and between neighbouring R groups, there is a limit to the permitted set of dihedral angles; the steric repulsion acts as a limit controlling on the distribution of conformations. This implies that the primary sequence of a polypeptide has a significant influence on the resulting higher order structures that forms.

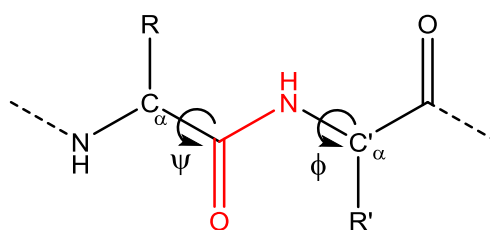


Figure 1.3: Structure of two generic amino acids as part of a peptide chain, with the dihedral angles displayed (ψ and ϕ) and the planar peptide bond in red.

The Ramachandran plot depicts this distribution of allowed conformations (Figure 1.4), as modelled by Ramachandran in 1963, using hard spheres with radii equal to their van der Waals radii to model atoms⁹. It is clear from this visualization that there are a limited number of permitted angle pairings, with the majority being sterically forbidden. Certain dihedral angle pairings being about distinct secondary structure motifs, with the driving force for stabilization of the resultant structure being predominantly hydrogen bonding.

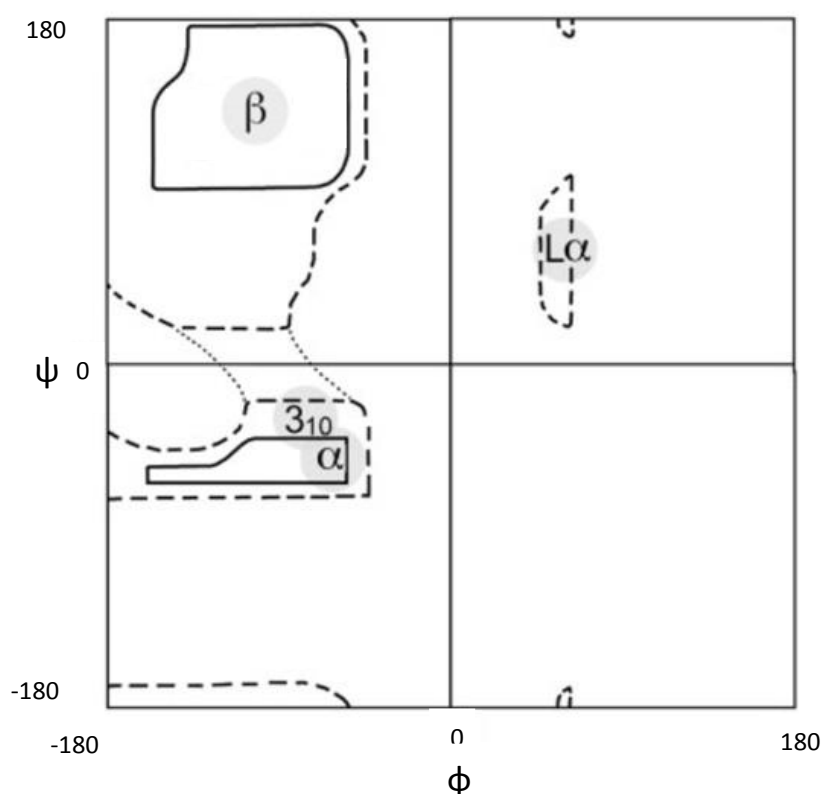


Figure 1.4: Ramachandran plot displaying the permitted dihedral angles (ψ and ϕ) for a polypeptide chain, along with areas that are associated with various secondary structures (left and right-handed α -helices, β -sheets, and 3_{10} -helices). Areas within the solid lines are favoured angles and areas between the solid and dashed lines are permitted but less favourable angles¹⁰ (Figure courtesy of Wikimedia commons user Dcrjsr released to the public domain).

The DSSP algorithm (Equation 1) is used to determine a hydrogen bond, which can only exist if the energy of the bond (E) is lower than -0.5 kcal/mol (-2.1 kJ/mol)¹¹. DSSP can be further used secondary structure motifs, dependent on the relative position and repetition of proximal hydrogen bonds, with each motif being given a single letter code, examples in Table 1.1.

$$E = q_1 q_2 \left\{ \frac{1}{r_{ON}} + \frac{1}{r_{CH}} - \frac{1}{r_{OH}} - \frac{1}{r_{CN}} \right\} \cdot 332 \text{ kcal/mol} \quad (1)$$

Equation 1: Where $q_1 q_2$ are partial charges of ± 0.42 e and ± 0.20 e to the carbonyl oxygen and amide hydrogen respectively, and the r_{AB} terms indicate the distance between atoms A and B, taken from the carbon (C) and oxygen (O) atoms of the C=O group and the nitrogen (N) and hydrogen (H) atoms of the NHR₂ group.

Table 1.1: A number of examples of the DSSP algorithm

Single letter code	Name of motif	Additional information
G	3_{10} -helix (3-turn helix)	Minimum length 3 residues Dihedral angles (ϕ , ψ) approx. (-49° , -26°)
H	α -helix (4-turn helix)	Minimum length 4 residues Dihedral angles (ϕ , ψ) approx. (-60° , -45°)
I	π -helix (5-turn helix)	Minimum length 5 residues Dihedral angles (ϕ , ψ) not well defined
E	parallel or anti-parallel β -sheet	Min length 2 residues Dihedral angles (ϕ , ψ) approx. (-135° , $+135^\circ$)
C	coil (residues which are not in a defined conformation)	

1.2.2.1 α -helices

The α -helix is one of the most common secondary structure motifs, manifesting itself as a regularly repeating right-handed helix with each of the R groups pointing outwards from the helix which reduces steric hindrance (Figure 1.5). Each constituent helix amino acid residue has the same set of dihedral angles (ϕ , ψ) (-60° , -45°), or more generally the ψ dihedral angle of one residue and the ϕ dihedral angle of the next residue sum to approximately -105° . These sets of dihedral angles give the helix a pitch of 5.4 \AA which is the vertical length of each repeating turn made up of 3.6 residues (a translation of 1.5 per residue relative to the previous residue). An α -helix is defined mainly through hydrogen bonding, the official definition nomenclature being $i+4 \rightarrow i$ hydrogen bonding. This describes the hydrogen bonding between the N-H group on a certain amino acid within the helix and the C=O group of an amino acid four residues earlier. This α -helix is not the only helical motif that can form; the more tightly wound 3_{10} -helices and the more loosely wound π -helices (Table 1.1), being defined as $i+3 \rightarrow i$ hydrogen bonding and $i+5 \rightarrow i$ hydrogen bonding, respectively.

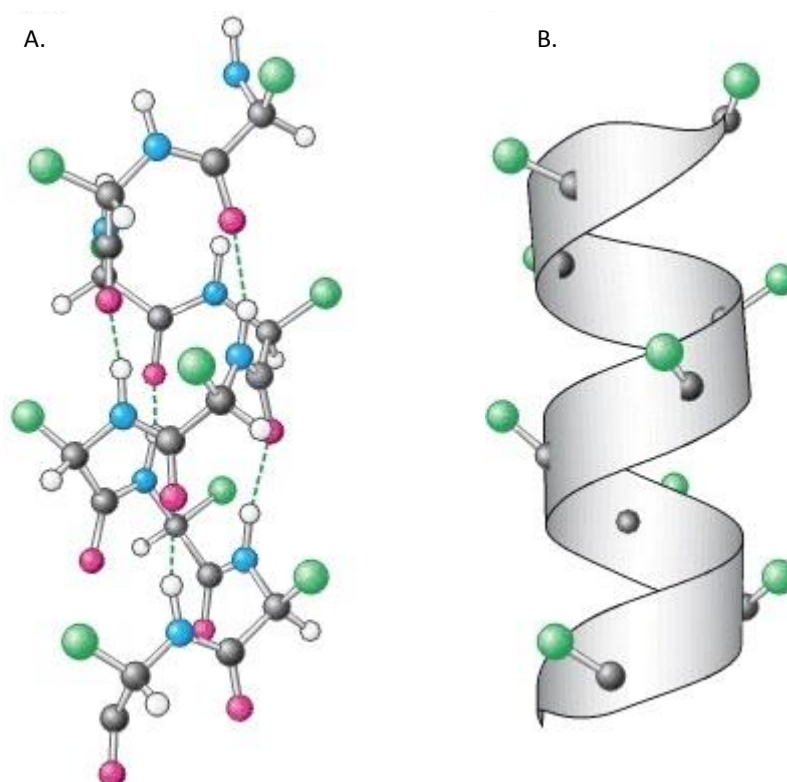


Figure 1.5: **A.** The side view of a ball-and-stick version of an α -helix depicting the hydrogen bonds (dashed lines) between **NH** and **CO** groups, **B.** Ribbon depiction with the α -carbon atoms and side chains (**green**) shown¹².

1.2.2.2 β -sheets

Another commonly occurring secondary structure within proteins are β -sheets, formed from segments of the polypeptide backbone (known as β -strands) that are laterally connected to adjacent β -strands by two or three hydrogen bonds. β -strands are directional ($N \rightarrow C$) in the same way polypeptide chains are with the C_α atoms in adjacent strands lining up (Figure 1.6). The hydrogen bonds are between the amide hydrogen on one β -strand and the carbonyl oxygen on the other, often instigating a slight twist along the strand and giving the pleated appearance. This configuration leads to the dihedral angles (ϕ , ψ) being $(-135^\circ, +135^\circ)$, and the R groups on the residues alternating between being above and below the plane of the β -strand. Adjacent β -strands can be arranged in two conformations of β -sheet; the more common anti-parallel β -sheet occurs when the $N \rightarrow C$ directions of adjacent β -strands are opposite allowing the hydrogen bonding to be planar (Figure 1.6), and the parallel β -sheet occurring

when the N→C directions of adjacent β -strands are the same which causes the hydrogen bonding to be non-planar (Figure 1.6).

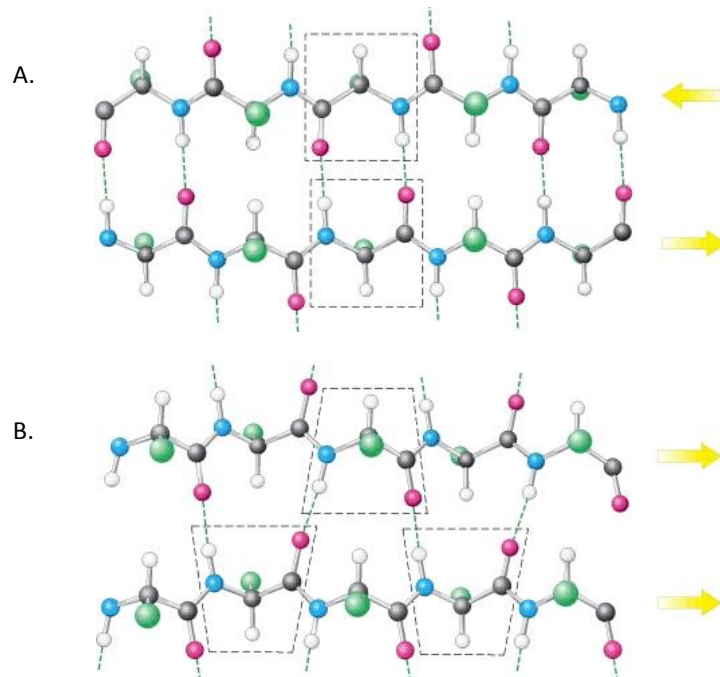


Figure 1.6: **A.** Antiparallel β -sheets: Formed of adjacent β -strands that run in opposing directions. Hydrogen bonding stabilises the structure, with NH and CO groups from one amino acid in the 'upper' β -strand forming two hydrogen bonds with one amino acid in the 'lower' β -strand. **B.** Parallel β -sheets: Formed of adjacent β -strands that run in the same direction. Hydrogen bonding stabilises the structure, with NH and CO groups from one amino acid in the 'upper' β -strand forming one hydrogen bond with two amino acids in the 'lower' β -strand.¹²

1.2.2.3 Other secondary structure motifs

Helices and pleated sheets are not the only stable secondary structures found within proteins, however, these often display a lower periodicity in comparison. The next most common motifs are generally classified as loops and turns, these are areas of secondary structure that are required to maintain the compact nature a protein possesses and are well defined in comparison to unordered/random coil segments of the protein. There are numerous other secondary structure motifs, often being combinations of the previously defined structures.

1.2.3 Tertiary structure

A protein's tertiary structure refers to the overall shape a single polypeptide adopts when fully folded, which is influenced by the residues in the primary sequence and the secondary structures that form due to the ordering of the primary sequence. The intramolecular interactions that occur between R groups on the residues and the hydrophobic effect determine the tertiary structure of a protein.

The function a protein intrinsically possesses is inherently linked to the tertiary structure that forms, with the 3D nature of the protein giving rise to active sites, often formed within crevices or open folds in solvent accessible areas of the protein. Enzymes use these spatially-constricted residue-specific

areas to increase the probability of favourable interactions with specific residues that are not necessarily neighbouring in the primary sequence. Proteases are enzymes that catalyse the hydrolysis of peptide bonds in substrate peptides/proteins, with specific active site residues used to bind the peptide to the active site and other residues to act as catalysts and cleave the peptide bond. If the protease was not folded into a specific 3D structure, these few residues (relative to the overall number) would not be located in these favourable positions, implying that the function is only possible, and is a consequence, of the folded tertiary structure¹².

The tertiary structure of a protein, as well as being formed of intricately folded amino acid residues, also includes cofactors; a non-protein complex organic molecule or metallic ion that interacts with the protein and is often a prerequisite for protein activity. Tightly bound cofactors are called prosthetic groups, *e.g.* the haem group within haemoglobin and myoglobin¹³, or the inorganic ion with a metalloprotease¹⁴. The bound prosthetic group within a metallo-protease, being a metal ion, is present to provide interactions with substrates that the protein cannot perform itself, often good electron donor/acceptor (Figure 1.7). If the prosthetic group is not present or is not correctly bound to the protein, the protein is functionally inactive. Other cofactors termed co-substrates are transiently bound to precise locations on/within the protein, and once bound the protein is activated, *e.g.* the binding of a molecule could bring about the active form of the enzyme through a conformational change¹⁵. Once the co-substrate is no longer bound, either through an environmental alteration or the co-substrate being sacrificial, the protein reverts to its inactive state, *e.g.* the protein shifts to a non-active conformation.

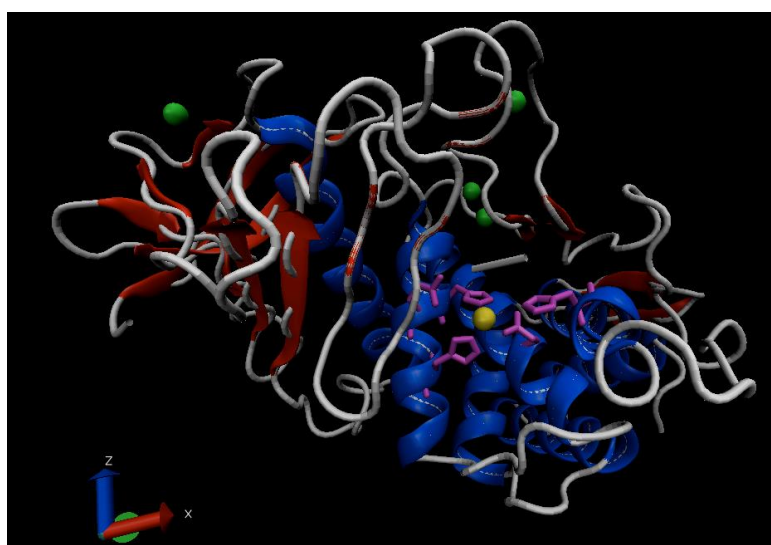


Figure 1.7: Ribbon model of a metalloendopeptidase (thermolysin), displaying the Alpha-helices in blue, beta-sheets in red, Zn²⁺ ion in yellow, active site side chains in magenta, Ca²⁺ ions in green¹⁶.

1.2.4 Quaternary structure

Not all proteins are formed from one single folded polypeptide chain, some more complex proteins utilise numerous folded polypeptide chains (subunits) that form a multi-subunit protein complex, the arrangement of which is the quaternary structure. The subunits in these protein complexes can be identical repeats (homodimer) or multiple different polypeptide chains (heterodimer). A classic multi-subunit protein is haemoglobin, being formed of four haem-containing globular protein subunits (the subunits being termed α - and β -chains) that are arranged in a clockwise α - β - α - β fashion (Figure 1.8). Haemoglobin has two conformational states; a deoxygenated state (termed T-state) and an oxygenated state (termed R-state)¹⁷. This indicates that the binding of an oxygen molecule to one of the subunits induces a change in the tertiary structure of the individual subunit. It is interesting to note that the binding of subsequent oxygen molecules requires a lower energy, implying that the conformational change in one subunit influences other subunits, and therefore the quaternary structure.

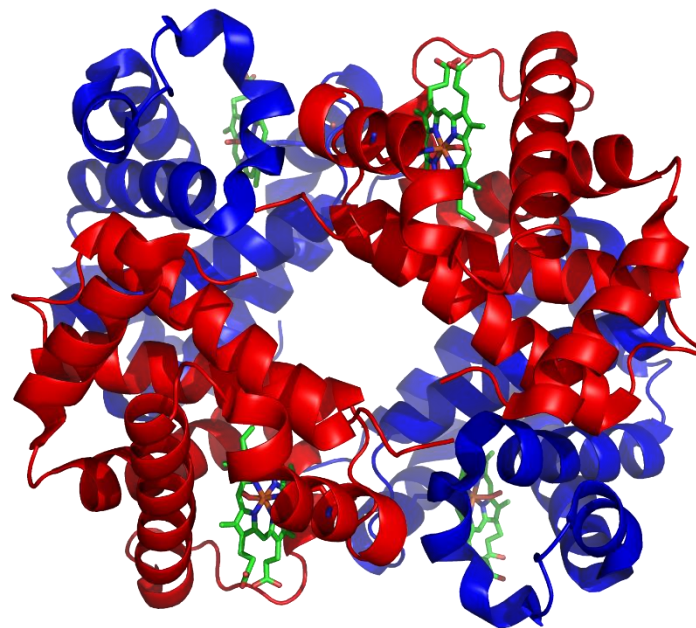


Figure 1.8: General ribbon structure of haemoglobin with the α and β units, in red and blue respectively, along with the iron-containing haem groups in green¹⁸.

1.2.5 Stabilization of protein structures

For a protein to fold and form highly ordered structures (whether they are secondary, tertiary, or quaternary) the free energy difference (ΔG_u) between the unfolded state (G_u) and the folded state (G_f) must be positive. This indicates that the folded state is more thermodynamically stable. The Gibbs free energy of the system is a balance between an enthalpy term (H) and an entropy term (S) related *via*

the Gibbs free energy equation (Equation 2). It is important to note that the Gibbs free energy of the system does not only factor in the entropy and enthalpy of the protein itself but also the of the solvent that surrounds and interacts with the protein in question. Entropy is the major energetic barrier obstructing the folding of a protein; the compact nature associated with the native state of a folded protein introduces high levels of order into the system that is not equal to the additional freedom that the solvent molecules experience¹⁹. Typically, the native state of a protein is more stable than the unfolded state by -5 to -15 kcal.mol⁻¹, which is low in comparison with the bond dissociation energy of a C-C bond at 350 kcal.mol⁻¹²⁰. This indicates there is a fine balance between the energy penalty from the entropic term and the enthalpic energy gain from the ensemble of forces that create numerous bonds throughout the folded protein.

$$G = H - T\Delta S \quad (2)$$

Equation 2: Where G is the Gibbs free energy, H is the enthalpy, S is the entropy, and T is the temperature in Kelvin²⁰.

1.2.5.1 Hydrogen bonding

Hydrogen bonding between the amide hydrogens and the carbonyl carbons is the primary force involved in the formation and subsequent stabilization of secondary structure motifs such as α -helices and β -sheets, as detailed previously. However, these potential hydrogen bonds (under aqueous conditions) can be equally satisfied through hydrogen bonding with the solvating water molecules. As such, only taking hydrogen bonding into account, the protein is no more energetically favourable in the folded or unfolded state. Therefore, whilst hydrogen bonding is evidently a significant factor in the stabilization of a protein, other forces need to be considered to understand how a protein combats the large decrease in entropy upon folding.

1.2.5.2 The hydrophobic effect

A dominant driving force for the folding and stabilization of a protein is the removal of water molecules from the hydrophobic (non-polar) residue environment through the process of burying them within the protein. Proteins have evolved to contain the majority of the hydrophobic residues within its core, shielding them from non-favourable residue/water interactions and increasing the more favourable water-water interactions²¹. The burying of these residues also increases the entropy of the water molecules, as they have been shown to form structured cages around hydrophobic residues when solvating. It has been calculated that the hydrophobic effect accounts for a large percentage of stabilization energy, indicating the importance of the hydrophobic effect for the retention of a native

structure²². However, smaller globular proteins cannot completely bury all the hydrophobic residues, therefore even though other forces contribute only a small percentage to the overall stabilization, they are not to be trivialized.

1.2.5.3 *Van der Waals interactions*

The high-density packing of the hydrophobic core and the close packing of the secondary structure motifs found within a folded protein leads to a vast number of van der Waals interactions²³. Whilst individually being small in stabilization energy, the high number of van der Waals interactions displayed throughout the protein add up to a non-trivial contribution.

1.2.5.4 *Salt bridges, disulfide bonds, and ionic bonds*

The previously detailed stabilization effects (hydrogen bonding, the hydrophobic effect, and van der Waals interactions) are all non-specific. There are also a number of specific stabilization effects that only occur between particular groups of amino acid residues, due to the R group presented. These effects have been shown to draw distant areas of proteins together, increasing the overall packing density. The first of these is electrostatic interactions or the formation of ionic bonds. This occurs between a positively charged amine residue (lysine, arginine, and occasionally histidine) and a negatively charged carboxylic acid residue (glutamic acid and aspartic acid). A second is a combination of electrostatic interaction and hydrogen bonding termed a salt bridge (Figure 1.9A). This effect occurs often between two α -helices when a lysine residue and a glutamic acid residue interact increasing the strength of the hydrogen bond formed, and further increasing the stabilization of the tertiary structure. A third stabilizing contribution is the formation of a disulfide bond between two local cysteine residues through the oxidation of thiol moieties (Figure 1.9B)²⁴. The stabilization due to this effect is large as the bond formed is a covalent S-S single bond, however, there are often only a few disulfide bonds formed across a protein. Disruption of disulfide bonds is known to have drastic effects on the conformation and consequently the function of the protein.

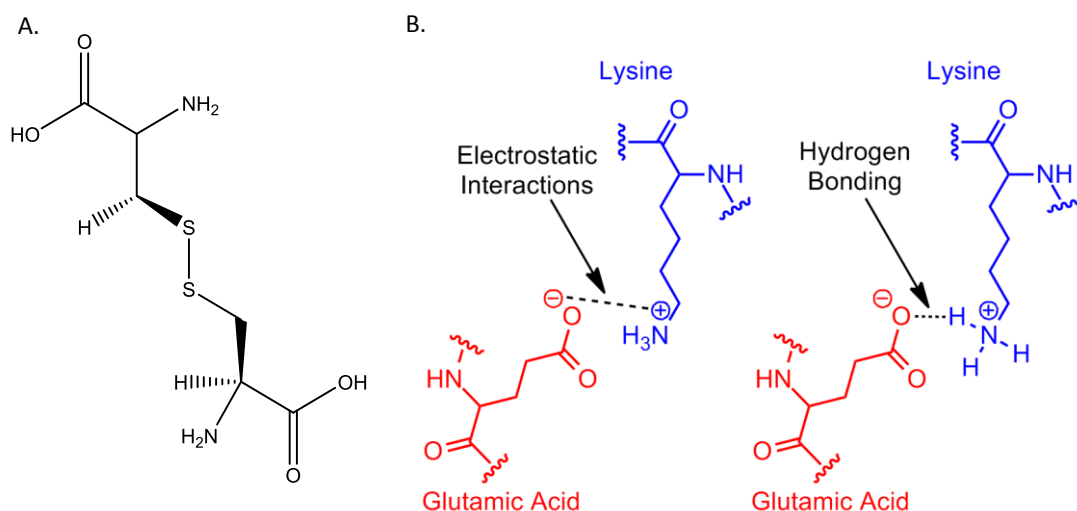


Figure 1.9: **A.** An example of a disulfide bond between two cysteine residues, and **B.** an example of a salt bridge between a glutamic acid and a lysine residue

1.2.5.5 Misfolding leading to diseases

The majority of proteins have well-defined folding pathways, often involving the formation of transient intermediate states of the protein. These intermediate states represent local minima in the folding pathway, with the global minimum being the native folded state. The local protein environment is key for the protein to follow the most energy efficient folding pathway, with perturbations to the ideal environment often inducing stable misfolds in the structure or stabilization of energetically trapped intermediate species, rendering the protein mis- or non-functional. Cells within the human body (as well as most other living organisms) have so-called quality control mechanisms that are proficient at the detection, localization, and removal of incorrectly folded proteins, however, they are not 100% effective. Misfolded proteins that are not removed will frequently aggregate due to the now exposed hydrophobic residues favouring interaction with similar non-polar residues rather than the polar nature of water. The misfolding of proteins and potential aggregation of such proteins can lead to a number of genetic diseases such as cystic fibrosis and numerous neurodegenerative diseases such as Alzheimer's disease²⁵. Mutations in genes that encode the cystic fibrosis transmembrane conductance regulator (CFTR) protein lead to misfolding and blockage of the channel preventing the flow of water and Cl⁻ ions, causing build-ups of thick mucus in several areas of the body, leading to eventual infection and blockages²⁶.

1.3 The effect of modification and bio-conjugation on the structure and function of proteins

As detailed in the previous sections, proteins are a set of bio-functional molecules with an array of simple outwardly displayed functional groups, and yet they can display a wildly varying set of functions that can include; energy storage and transfer, chemical signalling, catalytic and anti-pathogen capabilities, and structural support (e.g., actin)²⁷. However, the 20 different amino acid residues

display a limited set of functional groups. Therefore, as the body contains an extensive assortment of proteins that display functional groups and other molecules not found in this group of amino acids it must alter them post-synthesis. Accordingly, described below are detailed a number of post-translational modifications that can occur within cells and how they affect the function and structure of the protein.

1.3.1 Post-translational modifications

Post-translational modifications (PTMs) are modifications that occur to proteins in human cells and in cells of many other living beings. This refers to the modification of a protein after it has been synthesised (translated from mRNA to a polypeptide chain) through the covalent linking of functional groups or cleavages of peptide segments²⁸. These PTMs are frequently conducted through the use of enzymatic interactions and often alter the configuration and function of the protein. PTMs are a simple method for cells to increase the number of functions one protein can have and are crucial for correct cellular function due to many intracellular and extracellular events requiring specifically configured proteins or proteins displaying specific functional groups.

1.3.1.1 *Charge altering modifications*

1.3.1.1.1 Phosphorylation

Phosphorylation is a PTM to a protein that involves the covalent addition of a charged phosphoryl group (PO_3^{2-})²⁹ to an amino acid side chain and can occur to the majority of residues. The phosphorylation of proteins is performed by kinases: the transfer of the γ -phosphate group from high energy adenosine triphosphate (ATP) to a lower energy protein substrate displaying a hydroxyl moiety. The addition of a phosphate molecule to a hydroxyl group within a hydrophobic region of a protein, mediated *via* a kinase, can convert it to a hydrophilic region. This, due to the interaction with neighbouring residues, initiates a conformational change which can induce the start of signalling pathways within receptors and activate/deactivate enzymes, *e.g.* kinase. The basal state of the kinase is dephosphorylated and inactive, with activation occurring through the phosphorylation of the kinase *via* another kinase. The phosphorylation induces a conformational change within the active site, leaving the kinase in the active form.

1.3.1.1.2 Sulfation

Sulfation is a PTM to a protein that involves the covalent addition of a charged sulfate group (SO_4^{2-})²⁹ to an amino acid side chain, often occurring to a tyrosine side chain within the trans-Golgi and catalysed *via* the tyrosylprotein sulfotransferase protein³⁰. As with phosphorylation, the addition of a highly charged molecule to a site on the surface of the protein is known to alter the structure of the protein, potentially changing its function. Sulfated tyrosines have been shown to form stable salt

bridges between local proteins, increasing protein-protein interaction that can be critical for correct biological activity to occur. The covalent binding of two proteins alters the secondary and tertiary structure of both proteins in question, not only due to the bond formed but through the increase of additional protein-protein interactions that will take place due to the spatial proximity of the proteins.

1.3.1.1.3 Lipidation

Lipidation is a PTM of a protein that involves the covalent addition of one or more lipids, *e.g.* fatty acids, to a residue displayed on the exterior of the protein; the specific surface residue being lipid-head dependant³¹. The central purpose of the bound lipids is to anchor the protein to a specific membrane, *e.g.* endoplasmic reticulum, with the target membrane being lipid-tail dependant. Once the lipid tail is embedded with the selected membrane alongside similar fatty acid tails, the protein is known as a lipid-anchored protein or membrane protein.

There are two generalised types of lipidation, varying on the location; one occurring within or upon the surface of the cytoplasm membrane (N-myristoylation) and the other occurring within the secretory pathway, specifically within the lumen (glycosylphosphatidylinositol (GPI)). N-myristoylation is the covalent linking, *via* an amide linkage, of myristate, a fatty acid, to a glycine residue located at the N-terminus³². GPI anchors tether proteins to the plasma membrane, also allowing the transport of the protein from the ER through the secretory pathway to the extracellular face of the membrane³³.

A number of membrane proteins require lipidation to fold correctly, with the lack of lipid binding causing misfolding in the structure and the incorrect orientation or helices, altering the functionality of the protein. Post-translational addition of lipids to some proteins results in the folding of the protein to the native conformation due to the interaction of the protein with the lipid either through associated hydrogen bonding or through the lipid being charged. The correct folding leads to activation of the function of the protein³⁴. This is especially true of membrane proteins, known to lose structural integrity along with the correct membrane function when lipids are not present.

1.3.2 Peptide cleavage

Peptide cleavage can be a PTM that involves the hydrolysis of a peptide bond by a protease, often removing an initiator residue (methionine) or a small sequence of amino acids (generally requiring two cuts). The cleaving of a peptide bond can activate/mature a protein (insulin and renin), the cleaving can remove auto-inhibitory domains (sigma⁷⁰)³⁵, and also detach signal peptides from native proteins (Pep²⁷)³⁶.

Cleaving a peptide bond can drastically change the secondary and tertiary structure of a protein. For example, as part of a PTM of proinsulin, the protein synthesised prior to becoming insulin, it is cleaved in two places by an endopeptidase (prohormone convertases) removing a 35 residue long peptide sequence from the centre of the protein³⁷. This leaves behind two peptide chains that are bound by

two disulfide bonds, a drastically different conformation to the synthesised protein that can now bind to insulin receptors³⁸. Human prorenin is another example, with a 46 residue long peptide sequence being cleaved to produce renin³⁹.

Many enzymes are also post-translationally modified in the form of a peptide bond cleavage. Enzymes are generally translated in a proenzyme or zymogen form, often due to the enzyme being potentially harmful to the surrounding environment; e.g. pepsin is translated in the form of pepsinogen, with early activation of pepsinogen leading to acute pancreatitis⁴⁰. Before the zymogen can be an active enzyme it must undergo a PTM, often involving peptide cleavage that leads to the refolding of a specific length of the peptide chain. This refolding either alters the conformation of the active site or reveals the active site making it solvent accessible, thus activating the enzyme.

1.3.3 Bio-conjugation

Bio-conjugation refers to the covalent derivatisation of biomolecules⁴¹, i.e. the chemical modification of biological molecules to alter the function or environmental interaction of such molecules. It is imperative to have a large range of protein modification techniques along with the understanding of how these modifications affect a protein as many proteins, specifically enzymes, are widely utilised within many chemical procedures in a vast array of scientific disciplines.

Throughout the latter half of the 20thC, a number of protein-specific techniques were created and investigated, e.g., site-specific and side-chain modifications, crosslinking, and lipidation are all furthering the field of protein chemical modification. The historic invention of site-directed mutagenesis, a method used to intentionally alter specific regions of a gene's DNA sequence in order to investigate structure and activity of a protein, was also discovered in this time-period⁴².

1.3.3.1 Enzyme/substrate modification

Enzyme function requires an enzyme to be in a specific conformation, with the correct residues within the active site displaying the enzymatically necessary functional groups. How the enzyme is folded determines the size and shape of the active site, and therefore the substrates that it can cleave. An incorrectly folded enzyme may have an altered active site leading to potential steric hindrance and the prevention of substrate entrance. Similarly, if the correct residues within the active site are not present, the enzyme cannot perform the desired function⁴³. This applies specifically to residues that are members of the catalytic diad/triad, or that are required for the enzyme to interact with the substrate, i.e. the substrate binding cleft residues or those involved with the oxyanion hole (Figure 1.10).

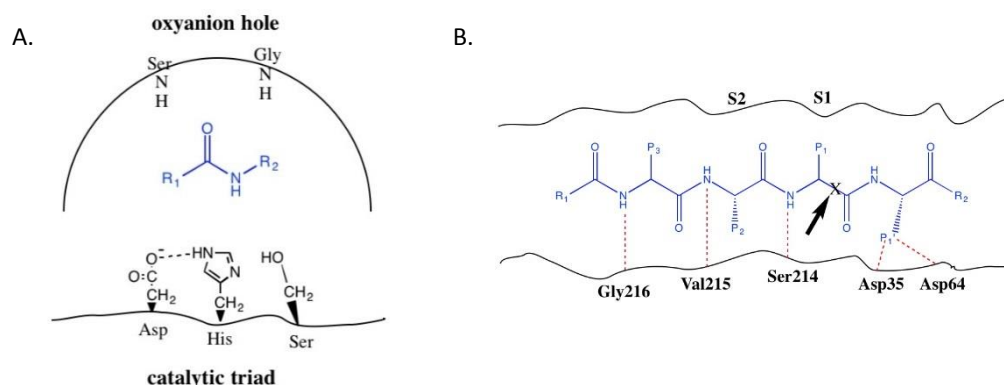


Figure 1.10: Cartoons displaying examples of **A.** An oxyanion hole (Ser and Gly) and a catalytic triad (Asp, His, and Ser), and **B.** An example of a substrate binding cleft with the substrate recognition sites S1 and S2 aligning with complimentary sites P1 and P2 on the substrate molecule to cleave site X. Amino acids (black) and the substrate molecule (blue) are present in both.⁴⁴

As previously discussed, the chemical surface-modification of a protein through the covalent binding of another molecule to the protein surface alters the displayed functional groups and has the potential to alter the folded conformation. Thus, the chemical modification of an enzyme, be that a surface modification or one within the active site, can influence the function of the enzyme. Nevertheless, the chemical surface-modification of an enzyme can be beneficial in numerous ways.

1.3.3.2 Protein stability

One of the major challenges faced when implementing the practical use of enzymes as biocatalysts in complex reaction syntheses is the stabilisation of the enzyme during storage and operation, *i.e.* organic solvents along with high temperature and pressure environments⁴⁵. It is imperative that the enzyme is stabilised without the diminution or dilution of activity; these being counterproductive to their utility. The stability of an enzyme comes in two distinct varieties that form the two stages of protein unfolding; thermodynamic and kinetic stability, Equation 3⁴⁶. Thermodynamic stability relates to how easily a protein is unfolded to a reversibly denatured and in-active conformation, whereas kinetic stability relates to how easily a protein enters an irreversibly denatured conformation, often involving aggregation or covalent alterations to the protein⁴⁷. Work on protein stability assumes an established positive-correlation between structure and specific functionality, *i.e.* if the protein retains the specific functionality it therefore meets the protein-specific minimum structural requirement. In general, if a protein has a structural similarity lower than 100% when compared to native, then the catalytic rate can be higher but the specificity of the protein's activity can be lower than that of the native⁴⁸.

$$N \rightleftharpoons U \rightleftharpoons I \quad (3)$$

Equation 3: Where N is the native conformation, U is the reversibly denatured enzymatically in-active state, and I is the irreversibly denatured enzymatically in-active state. With $N \rightarrow U$ representing thermodynamic stability and $N \rightarrow I$ representing kinetic stability⁴⁷.

1.3.3.3 Thermal stability

One major issue with the stabilisation of enzymes is protecting them against thermal-denaturation; the introduction of thermal energy to the system that disrupts hydrogen bonds and salt bridges within a folded enzyme leading to denaturation and a loss of function. A group of proteins, including some enzymes, known as thermophiles, are more resistant to thermal-denaturation than other enzymes⁴⁹. There are several hypothesised explanations for their increased thermal stability when compared to mesophilic proteins. Thermophilic proteins exhibit improved internal packing and higher entropic stabilisation, having a lower entropy gain on folding implying retained residual structure when in the denatured state⁵⁰. Fewer hydrophobic residues are displayed externally, with a larger majority being packed internally, removing their energetically unfavourable interactions with water molecules⁵¹. Thermophiles also have a higher number of hydrogen bonds, disulfide bonds, and salt bridges distributed throughout the enzyme, increasing the stability of the folded structure⁴⁹. Many successful site-directed mutagenesis experiments have been conducted, converting externally displayed hydrophobic residues for hydrophilic residues, or the conversion of residues to ones that more favourably hydrogen bond. However, site-directed mutagenesis of residues within a protein can be a complex procedure with a high potential for disruption to the secondary structure and function of the protein.

One bioconjugation-based method for thermally-stabilising an enzyme would be to increase the number of intra-molecular or inter-subunit covalent connections, mimicking the higher number of salt-bridges and hydrogen bonds *etc.* present in thermophilic proteins⁴⁷. This method is known as crosslinking; the process that links two polymer chains through the formation of covalent bonds, be that between two proteins (intermolecular) or two regions of a single protein (intramolecular), which are formed *via* chemical reactions between the reactive groups of crosslinkers and the functional groups externally displayed on the surface of the protein⁵². This increases the energy required to denature the protein due to the increased number of bonds. Crosslinking is carried out using a variety of short-chained molecules displaying reactive functional groups, *e.g.* glutaraldehyde and N-Hydroxysuccinimide-ester⁵³. However, crosslinking a protein does not only make it more thermally stable, it can also increase the kinetic stability and the stability within organic solvents, both being important for industrial usage.

1.3.3.4 Stability within non-aqueous solvents

Correctly stabilized biocatalysts (enzymes) in near-anhydrous or completely-anhydrous environments are becoming increasingly sort after for industrial processes; the advantages being an increased output at a potentially lower cost, but also within industrially advantageous solvents^{54,55}. Whilst the high temperatures used in many industrial reactions can be mitigated through use of enzymes, the use of organic solvents or ionic liquids can be detrimental to the native enzymes⁵⁵. These non-aqueous solvents are often required for the solvation of the reactants in the synthesis, therefore, changing the solvent to water is not a viable option.

As detailed in Section 1.2, enzymes have evolved to fold into functionally relevant conformations when solvated within water, the protein-water interactions permitting the dynamical freedom required for protein activity, *e.g.* enzymatic catalysis. Enzymes that are solvated within non-aqueous solvents, such as organic solvents or pure ionic liquids, generally denature due to the lack of similar protein-water interactions⁵⁶. Some organic solvents often favourably interact with the hydrophobic core residues of the enzyme stabilising the non-active form of the enzyme and can also attenuate the dynamics leading to greatly reduced enzymatic activity.

Ionic liquids are organic salts that typically have a melting temperature below 100 °C and are being increasing used as solvents within industrially relevant reactions due to the high tunability of the chemical and physical properties^{57,58}. In the past 20 years the use of partially-anhydrous ionic liquids as an alternative solvent for proteins has increased^{56,59}. The interplay between the “hydrophobic/hydrophilic/amphiphilic character” of the ionic liquid ions and the water creates a stable environment for the “non-polar, polar, and charged surface areas” of the protein’s surface, with the protein retaining the correct conformation for functionality⁵⁶.

A stability versus solubility problem arises when a protein is solvated by organic solvents. Non-polar solvents are required for numerous processes as the reagents/substrates are often non-polar in nature and therefore do not dissolve well in water. Due to the majority of proteins evolving in water-based environments, the hydrophobic residues are mainly buried within the core of the protein, shielding the energetically unfavourable interactions between such residues and water molecules. Residues displaying polar functional groups, being hydrophilic, are present on the exterior of the protein, with the hydrogen bonds formed between such residues and the water molecules being crucial in retaining the conformation of the protein. When within a non-polar solvents, such as ethanol, these hydrogen bonds are not formed, and along with it being more enthalpically favourable for the protein to display the hydrophobic residues, it becomes more energetically favourable for the protein to unfold⁶⁰. The now unfavourable interaction of the hydrophilic residues with the non-polar solvent causes aggregation of unfolded proteins within the solvent, reducing the interaction of such residues with the

solvent. This therefore means that enzymes within organic solvents often have a low rate of enzyme activity due to alterations in conformation leading to reduced substrate accessibility to the active site.

When proteins are solvated within highly hydrophobic non-polar solvents, such as cyclohexane, interestingly the propensity to remain folded increases^{19, 61}. This is due to the peptide groups having favourable van der Waals interactions with cyclohexane, although no hydrogen bonding, and the non-polar side chains that are displayed externally strongly interacting with the cyclohexane relative to their interaction with water. However, the solubility of proteins within these highly hydrophobic solvents is extremely low, with mass aggregation and often the formation of two phases, with no enzyme activity present. Taking the polarity of solvent to the extreme, *i.e.* a vacuum, there is no solubility and the protein remains as a highly stabilized dry powder.

To reduce the aggregation and aid in the reduction of protein unfolding within organic solvents, proteins can be covalently or non-covalently immobilised onto supports; insoluble inert compounds such as silica or calcium alginate⁶². An 'immobilised protein' refers to "proteins physically confined or localized in a certain defined region of space with retention of any functionality, and which can be used repeatedly and continuously"⁶³. Being bound to a support through single or multiple points of attachments introduces additional bonds, be that *via* covalent or non-covalent interaction, to the protein that increases the protein's rigidity through a reduction in structural flexibility⁶⁴. Being more rigid increases the stability within organic solvents by increasing the energy barrier required to be surpassed by the hydrophobic effect in-order to unfold the protein. There is also a reduction in protein-protein interactions that lead to aggregation within organic solvents. Being attached to a large micro-sized object also makes the enzyme more readily extractable from a reaction, allowing the immobilised enzymes to be repeatedly reused, making the use of enzymes as biocatalysts more viable.

Crosslinking and immobilising enzymes can have a detrimental effect, often altering the surface charge and subsequently the interactions between hydrophobic and hydrophilic regions upon the surface. Changing these interactions has been shown in multiple cases^{41, 52, 65} to disrupt the tertiary structure of the protein and retard the catalytic activity, frequently rendering the protein inactive. However, a number of crosslinkers have been found to not alter the conformation, and correspondingly the activity, to too large a degree⁶⁶. As such, the crosslinking and immobilisation of enzymes has become routine and is widely utilised in areas such as mass spectrometry and biomedicine⁶⁶.

The covalent or non-covalent addition of surfactant molecules to the surface of enzymes can also have the effect of increasing stability⁶⁷. Surfactant molecular weight, chain composition, and active head/tail groups are many and varied, meaning that they can be utilised to stabilise enzymes in numerous environments. Polyethylene glycol (PEG) is a surfactant extensively used in industrial and pharmacological processes as it is biocompatible. PEGylation, the covalent linking of PEG with another molecule such as an enzyme, has the effect of cloaking the surface and increasing the molecular size. The physio-chemical properties of PEG are adopted by the enzyme as the PEG molecules are outwardly

displayed, although it has been demonstrated that enzymatic activity can be retained⁶⁸. As PEG is a water-soluble polymer, water-insoluble enzymes can be solubilised rather than precipitating or aggregating out of solution. Along with stability benefits, PEGylation or surfactant addition can impart other properties to the enzyme depending on the other moieties within the chain; improved thermal stability, increased enzymatic activity⁶⁹, and retained conformation.

A recent approach by Perriman *et al.*^{3, 70, 71, 72} has utilised these additional properties in order to form a solvent-free liquid protein, which has displayed a number of unique protein properties, including high-temperature stability and the retention of biological activity in this unusual phase of biological matter.

1.4 Solvent-free liquid proteins

It has been well documented that nano-scale objects, in comparison to molecular-sized objects, have a deep and narrow interaction potential due to the relative size of the particle to the interaction distance, meaning that the forces are strong at short distances. This relative size to interaction distance has a large impact on the phase behaviour of nano-scale objects, causing them to sublime or degrade when heated from the solid phase, as displayed in C_{60} ⁷³ and proteins⁷⁴. These physical changes in phase when heated signify that the liquid phase cannot easily be accessed. However, it was hypothesised that through increasing the range of these intermolecular interactions, the liquid phase of nano-scale objects can be accessed (Figure 1.11)⁷⁵. A crucial step forward was made in this field by Giannelis *et al.* when they functionalized a nanoparticle core with a 'covalently tethered ionic corona' creating a solvent-free nanoparticle liquid and demonstrated that it displayed similar properties to 'free-flowing liquid'⁷¹.

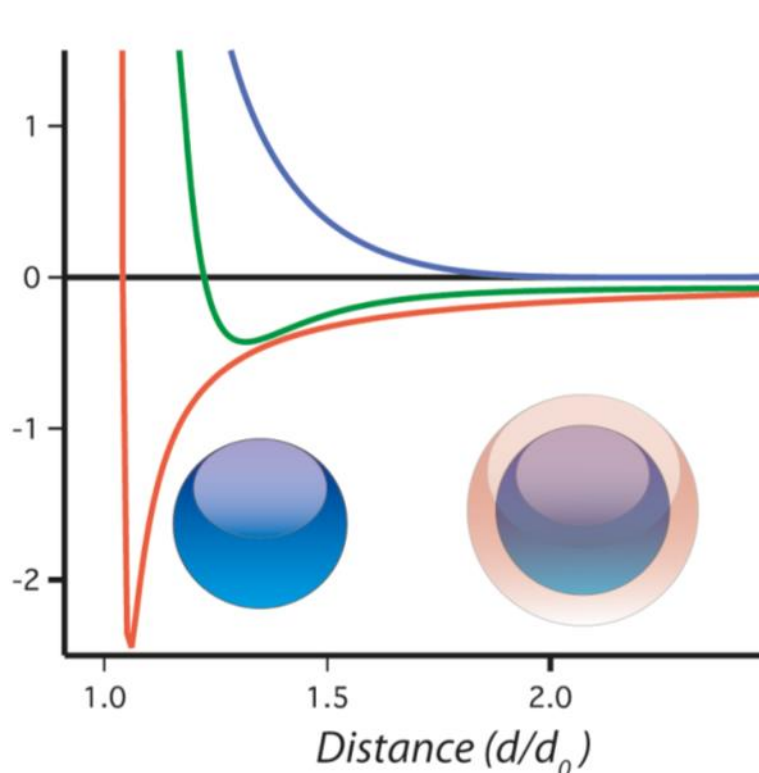


Figure 1.11: A graph depicting the deep and narrow attractive interaction potential well of a nano-scale object (orange line, blue sphere) alongside a sterically layered nano-scale object, the steric layer (orange) resulting in an increased repulsive potential (blue line) and therefore a more shallow and wide attractive potential well for the object (green line)⁷¹.

Proteins are nanoscale objects, with diameters in the range of 1-100nm, and therefore the interaction potentials experienced by proteins are similar to those of other nanoparticles. When proteins are heated from the solid state, rather than forming a liquid phase they thermally decompose, denaturing the protein and rendering it inactive⁷⁴. The absence of a protein liquid phase due to the short-ranged intermolecular interactions between neighbouring proteins relative to the size of the proteins, *i.e.* the interaction forces are too strong at short distances and too weak at distances of a few nanometers.

In a similar vein to Giannelis' constructs⁷¹, which are inorganic nanoparticles displaying a surfactant corona, Perriman *et al.* ionically coupled anionic surfactants to the exterior of a cationized protein, forming nanoconstructs with protein cores surrounded by a surfactant corona³ (Figure 1.11). Unlike the particles Giannelis developed, the protein constructs, created in an aqueous environment, could not be heated within a vacuum oven to remove the solvent and form a liquid as the hydration water surrounding proteins is very strongly bound. Instead, the sample was lyophilized to remove the water and then thermally annealed between 40 and 60 °C forming a solvent-free liquid protein (protein melt). The removal of the hydration water from a wild-type protein is known to alter or nullify the biological functionality of the protein due to the loss of solvent-protein interactions. Dehydration has been shown to constrict flexibility, leading to a reduction in motion and therefore reduced access to alternate conformational states⁷⁶, reducing/removing the biological activity. Despite this, range of biomacromolecules based solvent free melts are prepared and their physico-chemical properties have been investigated, each displaying retained secondary structures analogous to the native protein, along with retention of biological function^{3, 70, 77, 6, 127, 128}. Recently solvent-free protein liquid crystals have also been synthesised, with the LC behaviours being linked to the specific surfactant molecules used in the synthesis¹²⁹.

1.4.1 Synthesis of protein melts

Proteins, as previously detailed, are biomolecules consisting of one or more polypeptide chains folded into a precise conformation, often with bio-molecular activity/function. Each polypeptide chain is comprised of a specific arrangement of amino acid residues, some of which are charged and some of which that can, through pH adjustment, be protonated or deprotonated. The polypeptide chain is folded into a particular active conformation, with influence from the surrounding solvent⁷⁸, either self-folding or co-folding with other polypeptide chains. Due to the majority of proteins folding within an aqueous environment, the residues present on the surface of proteins have a higher propensity to be hydrophilic and possibly charged. These surface point-charges are known to be essential to the stabilization of the correct protein conformation and therefore altering this surface charge balance through; charge shielding, point mutations, cation- or anionization, can have a destabilizing effect, with the potential to alter the structure and denature the protein. The denaturation of a protein leads to a reduction in, if not total removal of, biological function, rendering the protein inactive⁷⁹.

1.4.1.1 Cationization

The initial stage in the synthesis of solvent-free liquid proteins is the cationization of the protein through the coupling of amine groups to the negatively charged surface residues; aspartic acid and glutamic acid, at a pH for the acidic groups to be deprotonated and the amino groups to be protonated. Cationization removes the negative charge from the acidic group, supplanting it with the positive

charge from the amino group, increasing the positive net surface charge of the protein by two per modification (Figure 1.12). The cationization is conducted through carbodiimide activation, cross-linking primary or secondary amines to the surface of the protein. This well-documented procedure has been shown to alter the secondary structure of the protein, nevertheless, proteins cationized *via* this method have displayed biological function⁶. This signifies the retention of an adequate percentage of secondary structure for the biological activity to remain, though generally in a reduced capacity relative to the native protein.

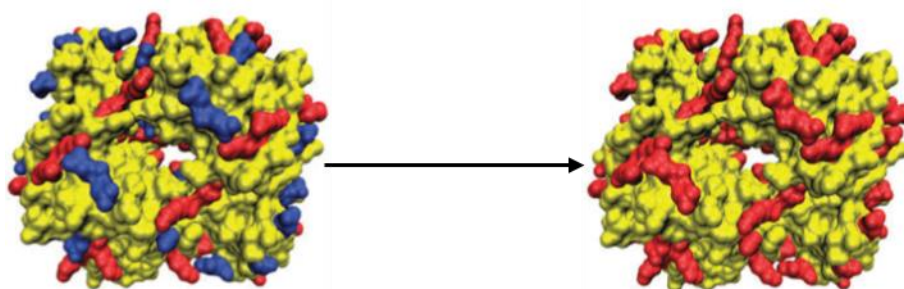


Figure 1.12: Schematic illustration showing the cationization of haemoglobin; negative charges (blue) and positive charges (red)⁸⁰.

1.4.1.2 Conjugation

The second stage in the synthesis of solvent-free protein melts is the addition of anionic surfactant molecules, a surfactant with a negatively charged head group, to the surface of the protein. These surfactants are electrostatically coupled to the cationized sites on the surface of the protein, with the opposite charges forming an ionic bond between the two molecules establishing a protein-surfactant nano-construct (Figure 1.13). A surfactant is an amphiphilic molecule, with each surfactant being formed of a hydrophobic tail group terminated by a hydrophilic head group. The charge on the head group is often the surfactant identifier, being cationic, anionic, amphoteric, or non-ionic, that being a positive, negative, both cationic and anionic, or uncharged head group respectively.

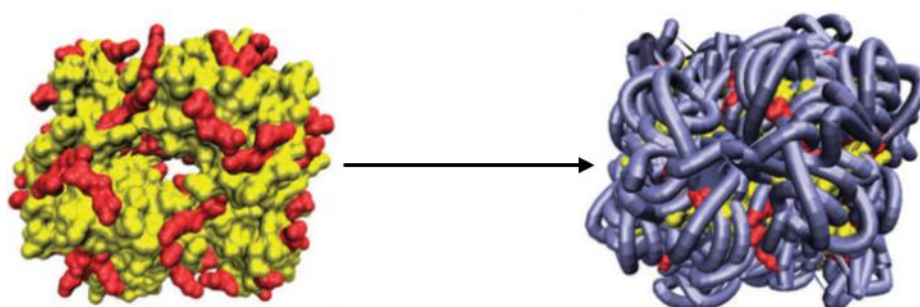


Figure 1.13: Schematic illustration showing the electrostatic coupling of an ionic-surfactant to cationized haemoglobin forming a haemoglobin-surfactant nano-construct; positive charges (red) and surfactant (purple tubes)⁸⁰.

Surfactants have long been used in conjunction with proteins, with the interactions between proteins and surfactants being broadly classified into two groups; those that denature a protein upon interaction and those that leave the protein conformation relatively unchanged upon interaction. Surfactants that display a long hydrophobic tail group, often having a lower hydrophilic-lipophilic-

balance (HLB), are often introduced to a protein in order to initiate unfolding of the polypeptide chain for an in-depth analysis of the chain structure. This unfolding occurs due to the hydrophobic tail interacting with the highly hydrophobic protein core, disrupting the present non-covalent bonds within the protein, and can remove all biological function. Sodium dodecyl sulfate protein polyacrylamide gel electrophoresis (SDS-PAGE) utilises the denaturation effect that the SDS has on a protein to determine the molecular weight of polypeptides⁸¹. Once the protein has begun unfolding, both anionic and cationic surfactants can then bind along the entire polypeptide chain, (Figure 1.14) at positively and negatively charged points respectively. Conversely, a surfactant displaying a long hydrophilic tail group, often having a high HLB, are often introduced to proteins to stabilise the protein, *e.g.* against unfolding or as aggregator preventers. The surfactant interacts strongly with the exterior groups on the surface of the protein, shielding them from interactions that could lead to denaturation and subsequent function deactivation.

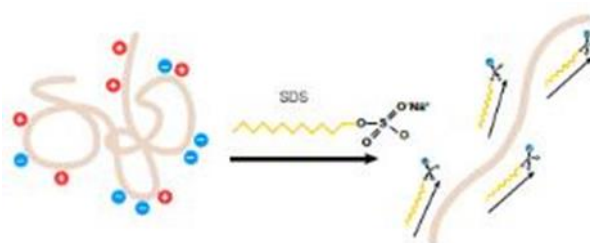


Figure 1.14: This cartoon displays how an anionically charged hydrophobic-dominant surfactant molecule can unfold and bind along the length of a polypeptide chain⁸².

Surfactants are often used in the capacity of protein stabilization within non-aqueous solvents, as previously detailed. The polarity of water is fundamental to the folding, conformational stability, and molecular dynamics of a protein. As such, removing a protein from an aqueous environment and placing it in an organic non-polar solvent can have a detrimental impact on the protein. Organic solvents often induce protein unfolding, increasing the interaction of the hydrophobic core residues with the solvent and reducing the external hydrophilic contact. This often provokes aggregation and a notable retardation in the catalytic rates of enzymes even if it remains folded as there is a reduction in functional conformations available⁴⁸. In this vein, surfactants have been employed in the dispersion of proteins within organic solvents, reducing the solvent accessibility to the protein. The hydrophobic tails of the surfactant also interact favourably with the solvent, aiding in the solvation of the protein within the relatively non-polar solvents. As the structure of the protein is somewhat retained with the utilization of surfactants, the catalytic rate of enzymes is often retained to a higher order⁸³. The use of surfactant-protein coupling has not only been utilised to improve the catalytic rates of proteins within organic solvents but also improve rates within water when compared to the native³⁸. This was theorised to be due to an increased substrate affinity attributed to the higher steric hindrance of the protein by the surfactant layer and the formation of rigid hydrogen bonds between surfactant and water molecules. The electrostatic coupling of a surfactant to a charge-modified protein has also been demonstrated to recuperate a substantial percentage of secondary structure⁷². This is hypothesised

to be due to the charge shielding effect that the oppositely charged surfactant head group has on the surface charge, *e.g.* +72mV to +7mV⁸⁰.

1.4.1.3 Protein films and solvent-free liquid proteins

As previously mentioned, the addition of surfactant to a protein forms a protein-surfactant nanoconstruct however, the rate at which the surfactant is added to the protein system affects the coating efficiency of the surfactant. Rapid addition of the surfactant, at a concentration above the critical micelle concentration (CMC), creates highly dispersed protein-surfactant conjugates. Slow addition of the surfactant, at a concentration below the CMC, initiates the spontaneous formation of protein-surfactant nanoclusters $\approx 100\text{nm}$ -diameter (Figure 1.15C). Once the nanoclusters have formed they arrange into a higher-order assembly creating $1\mu\text{m}$ -sized particles (Figure 1.15D)⁸⁴.

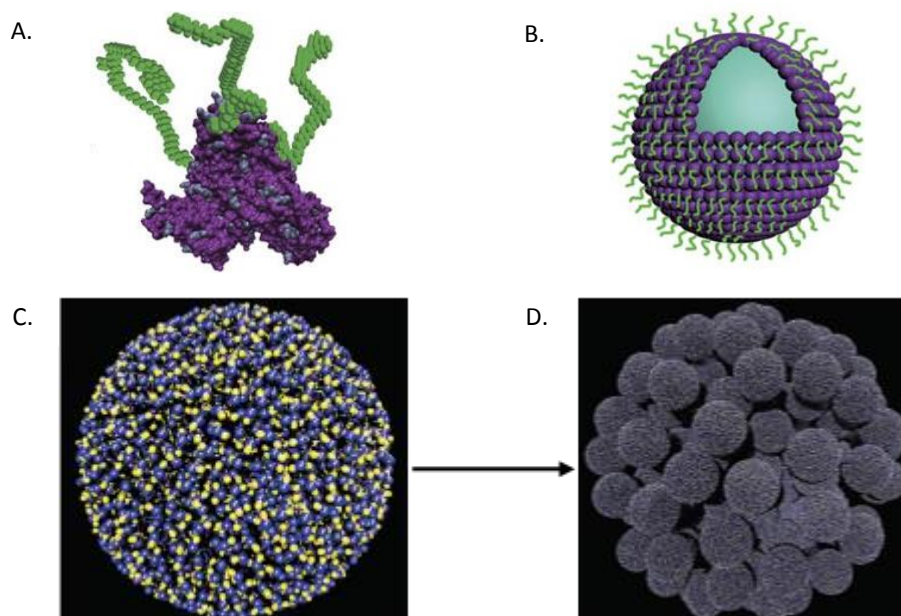


Figure 1.15: **A.** Slow addition of surfactant below the CMC of the surfactant causes the surfactant molecules to individually interact with the protein and creating a hydrophilic/hydrophobic imbalance over the surface of the protein creating a protein-surfactant nanoconjugate with a roughly 2:1 ratio of charged surface sites to oppositely charged surfactant molecules. **B.** The imbalance causing the spontaneous self-assembly of proteinosomes. **C.** These proteinosomes spontaneously form 100nm-diameter protein-surfactant nanocluster and **D.** a higher order assembly of nanoclusters creating a $1\mu\text{m}$ -sized particles⁸⁴.

In the case of slow addition, it is theorised that the addition of individual surfactant molecules to the surface of a protein partitions the protein into more hydrophobic and a less hydrophobic areas (Figure 1.15A) causing the proteins to aggregate into proteinosomes (Figure 1.15B). This aggregation effect terminates the surfactant coverage of the protein at a roughly 2:1 ratio of charged surface sites to oppositely charged surfactant molecules. The higher order assembly of nanoclusters can then be dried down and cross-linked forming self-standing protein-polymer surfactant films, with retained biological structure and function.

In the case of rapid addition, it is theorised that when a surfactant micelle interacts with a protein it envelops it, coating the entire surface with surfactant molecules in a 1:1 ratio of charged surface sites

to oppositely charged surfactant molecules (Figure 1.16A). Once lyophilized (Figure 1.16B) and then heated (Figure 1.16C) the conjugates interact with one-another through the surfactant tail domains at a length scale and strength that permits the formation of a liquid phase. This is in vast contrast to the thermal degrading that occurs when an un-modified dry lyophilized protein is heated, as the interactions between proteins are too strong and on too short a length scale. The formation of a solvent-free liquid protein also affords a number of unique behaviours to the protein.

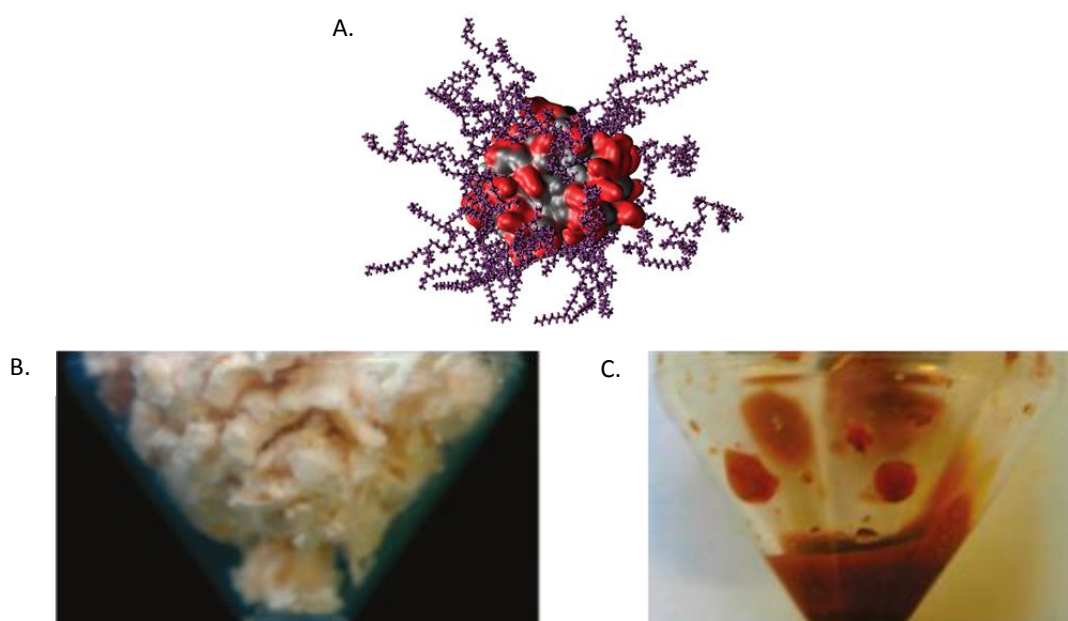


Figure 1.16: **A.** Rapid addition of surfactant above the CMC of the surfactant causes the surfactant micelle to envelop the protein and creates a protein-surfactant nanoconjugate with a 1:1 ratio of surfactant to positively charged surface site. **B.** Lyophilized [C-Ferritin][S] protein nano-conjugate and **C.** the same [C-Ferritin][S] sample post thermal annealing forming a solvent-free liquid protein³.

1.4.2 Physical characteristics of solvent-free liquid proteins

1.4.2.1 Phase behaviour of solvent-free liquid proteins

A number of physical and spectroscopic techniques; differential scanning calorimetry (DSC), rheometry, polarized optical microscopy (POM), small-angle X-ray scattering (SAXS) and wide-angle X-ray scattering (WAXS), have been used to analyse and characterise the phase behaviour of protein melts^{3, 84, 85}. Many of these techniques have displayed an indication of a viscoelastic protein liquid with liquid crystalline phase behaviour within a protein specific temperature range (32-37 °C for solvent-free liquid ferritin ([C-Fn][S])), along with evidence of a transformation to a Newtonian fluid phase above a protein specific temperature (40 °C for [C-Fn][S]). Significantly, DSC displayed evidence of a primary endothermic transition related to the melting of [C-Fn][S] occurring at 29 °C (Figure 1.17A) along with a discrete secondary endothermic transition at 36.5 °C for [C-Fn][S], indicating an intermediate phase. Rheology measurements at 50 °C displayed Newtonian fluid behaviour, however, once cooled to 32 °C the viscosity increased, with a marked increase in viscosity after a prolonged

period of time at this temperature (16 hrs). The gradual increase in viscosity demonstrates a slow temperature dependent re-ordering within the solvent-free liquid protein at this temperature. This gradual molecular reordering was further investigated using POM at 32 °C, with distinct Maltese crosses being observed throughout the solvent-free liquid protein. It was clear that the melt exhibited strong birefringence similar to a liquid crystal, with the Maltese crosses increasing in size when the melt remained at 32 °C for a prolonged period of time at this temperature (16 h). The increase in the size of the Maltese crosses indicates a larger liquid crystal domain size when the protein melt is cooled at a reduced rate. SAXS and WAXS data displayed short-range order up to 40 °C, leading to an isotropic scattering profile at 60 °C and evidence of alkyl tail-tail interactions and polyethylene glycol chain-chain interactions at 32 °C that are absent above 40 °C (Figure 1.17B)⁸⁵. Bragg peaks were also observed at 32 °C that relate to a mesolamellar phase with a (001) repeat distance of 13nm, similar to that of the external diameter of [C-Fn][S], with no evidence of the mesolamellar phase at 50 °C.

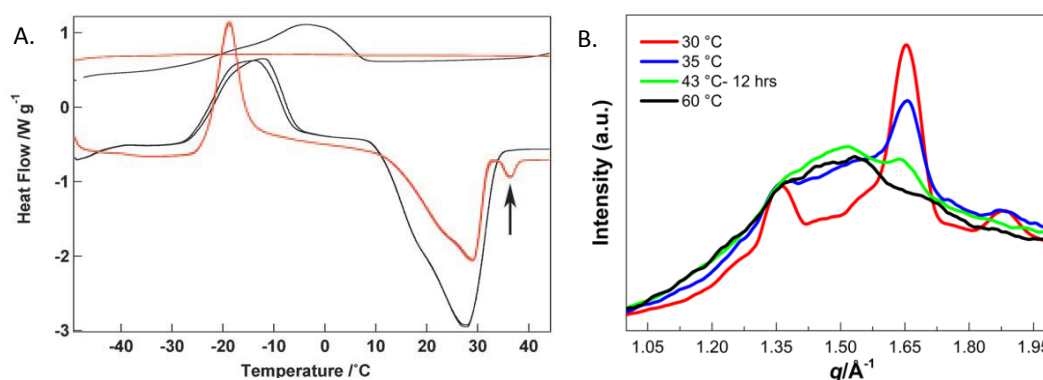


Figure 1.17: **A.** The DSC traces of anionic polymer surfactant S (black) and the [C-Fn][S] construct (red) displaying minima at 27.5 and 29 °C, being the melting transitions for the respective sample, the black arrow at 36.5 °C indicates an additional endothermic transition for the [C-Fn][S] construct. **B.** WAXS profile of solvent-free [C-GOx][S] at different temperatures showing two peaks at q values of 1.36 and 1.65 \AA^{-1} at 30 °C corresponding to polymer surfactant alkyl-alkyl or PEO-PEO correlation lengths, respectively, at 30 °C (red); a lower intensity predominant PEO-PEO correlation peak at a q value of 1.66 \AA^{-1} at 35 °C (blue), and in the solvent-free liquid held at 43 °C for 12 h (green); and an isotropic phase at 60 °C, indicating a rise in temperature decreases the amount of ordering⁸⁵.

All of these results indicated that the melt undergoes a slow kinetic-ordering transformation at 32 °C to a viscoelastic liquid with liquid crystalline behaviour and short-range ordering, transforming to a Newtonian-like fluid at 50 °C. The persistence of birefringence up to 37 °C and the loss of short-range structure between 35 °C and 43 °C confirmed the evidence from DSC, displaying a secondary transition at 36.5 °C. An in-depth structural and rheological analysis of solvent-free liquid glucose-oxidase [C-GOx][S] displayed analogous phase transitions, both in the primary melting and in the secondary viscoelastic transitions. However, the phase transition temperature ranges differed when [C-Fn][S] was compared to [C-GOx][S]. This is hypothesised to be due to [C-GOx][S] having a smaller hydrodynamic diameter ($d=9 \text{ nm}$) than that of [C-Fn][S] ($d=12.5 \text{ nm}$), along with fewer cationic surface charges once cationized. As a consequence, [C-GOx][S] has fewer surfactant chains with which neighbouring [C-GOx][S] surfactant chains can interact. The reduced number of surfactant-surfactant interactions alters the phase transition temperature as a lower thermal energy is required to break the electrostatic bonds.

1.4.2.2 Solvent-free liquid protein equilibrium structure and folding

Spectroscopic techniques; attenuated total reflectance-Fourier transform infrared spectroscopy (ATR-FTIR), Diffuse-reflectance ultra violet-visible spectrometry (DR-UV-Vis), and synchrotron radiation circular dichroism spectroscopy (SRCD), have shown that as well as interesting phase behaviours, solvent-free liquid proteins display not only relatively near-native secondary structure but also hyper-thermal stability and unique structural behaviour when heated and cooled^{70, 77, 72} through the use of ATR-FTIR it was shown that the solvent-free liquid protein, in this instance solvent-free liquid myoglobin [C-Mb][S], retained a near-native structure, with amide I and II band wavenumbers representing a protein with predominantly α -helical structure. This is in contrast to lyophilized native myoglobin that had bands representing a protein with increased random coil or β -sheet characteristics. This displays the benefit of the solvent-free liquid modification compared to a lyophilized native system, *i.e.* a higher retention of the native α -helical structure.

DR-UV-vis was utilized to study the Soret band found within haem-based proteins containing a haem prosthetic group; a shift in Soret, Q_α , and Q_β bands can indicate a change in haem environment. A marginally red-shifted Soret band and blue-shifted Q_α and Q_β , relative to aqueous Mb, for [C-Mb][S] was observed, with values “consistent with the presence of a six-coordinate, low-spin haem metalcentre with a strongly bound water molecule”⁷⁰. Through further investigations, it was established that the shift in band peaks was associated with the change in microenvironment polarity due to cationization rather than because of the removal of water molecules.

Further evidence for the retention of a large majority of the native secondary structure came from SRCD spectra, conducted at 25°C. From Table 1.2 it is evident that the cationization of myoglobin affects the secondary structure to a large degree, observing a large drop in α -helical structure and an increase in β -sheet, turns and unordered structure. However, it is also clear that the formation of [C-Mb][S] recuperates a significant percentage of the native structure, possibly due to charge shielding/neutralisation effects lowering protein-surface based repulsive interactions. When the peak intensity ratio, R (where $R = I_{208}/I_{222}$), of different structures, are compared it is again clear that the folded structure of [C-Mb][S] is comparable to native-Mb.

Table 1.2: Deconvoluted secondary structure content for various Mb constructs at 25°C, with S_2 representing neutralized ethylene glycol ethoxylate lauryl ether as the surfactant⁷⁰.

Sample	α -Helix/%	β -Sheet/%	Turns/%	Unordered/%
Solvent-free liquid [C-Mb][S_2]	51	9	11	29
Aqueous [C-Mb][S_2]	50	12	11	26
Aqueous C-Mb	42	14	13	31
Aqueous native Mb	81	2	6	11

Through temperature dependent SRCD and DR-UV-vis experimentation it has been shown that the thermal denaturation pathway of melt [C-Mb][S] is radically different to that of native-Mb. Significantly, the half denaturation temperature, T_m , of native-Mb is known to be 75 °C whereas melt [C-Mb][S] has a T_m of 160 °C⁷⁰. This displays a large increase in structural thermal stability due to the solvent-free liquid protein synthesis. Further evidence for this is that native-Mb is known to gradually reduce in secondary structure content as the temperature of the environment is raised, being completely denatured by 95 °C. In contrast, when [C-Mb][S] was heated to 95 °C, the α -helical content had only a 6% reduction (52%-46%), with the presence of α -helical structure up to a temperature of 235°C (Figure 1.18A), 140 °C higher than that of native-Mb. It is to be noted that at 95 °C aqueous [C-Mb][S] had lost 50% of the α -helical structure, displaying that whilst the surface modifications do have a large impact upon the thermal stability of a protein, the water removal stage in the protein melt synthesis increases the thermal stability of the construct further. A similar increase in structural thermal stability was evident from DR-UV-Vis (Figure 1.18B) displaying a limited shift in peak intensity and wavelength up to 100°C, indicating minimal structural perturbations, whilst native-Mb had fully denatured by this temperature.

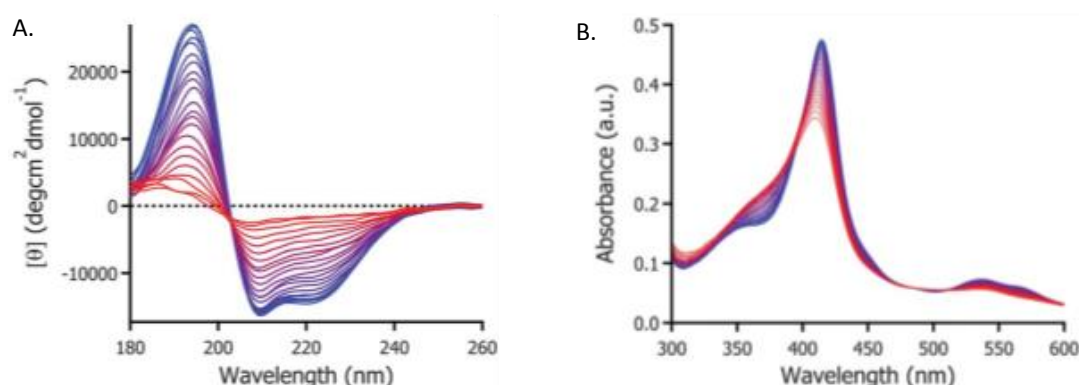


Figure 1.18: **A.** Synchrotron radiation CD (SRCD) spectra of solvent-free liquid [C-Mb][S] over a temperature range of 25 °C (blue) to 235 °C (red), showing the progressive reduction of peak intensities at 208, 222, and 195 nm at 10 °C intervals. **B.** DR-UV-vis spectra of [C-Mb][S] in the solvent-free state showing the progressive reduction of peak intensity along with a shift in the Soret band peak position, recorded between 25 °C (blue) and 175 °C (red) at 5 °C intervals⁷⁰.

Myoglobin, and other globular proteins, are known to refold to a near-native structure once heated to a temperature below the respective T_m and subsequently cooled to 30 °C. This structure recuperation was also observed in protein melts; [C-Mb][S] was heated to 155 °C and subsequently cooled to 30 °C, re-establishing 94% (49% from 52%) of the [C-Mb][S] α -helical structure (Table 1.3), and also recovering >95% of the Soret band (Figure 1.19). The highly efficient refolding of [C-Mb][S] displays retention of native thermal characteristics, albeit at higher temperatures, despite the surface modifications and lack of water molecules.

Table 1.3: Table of secondary structure content for solvent-free liquid [C-Mb][S] at 30 °C, 155 °C, and after cooling to 30°C⁷⁰.

Temperature/ °C	α -Helix /%	β -Sheet/ %	Turns/ %	Unordered/ %
30	52	7	12	28
155	38	15	13	33
30	49	10	11	30

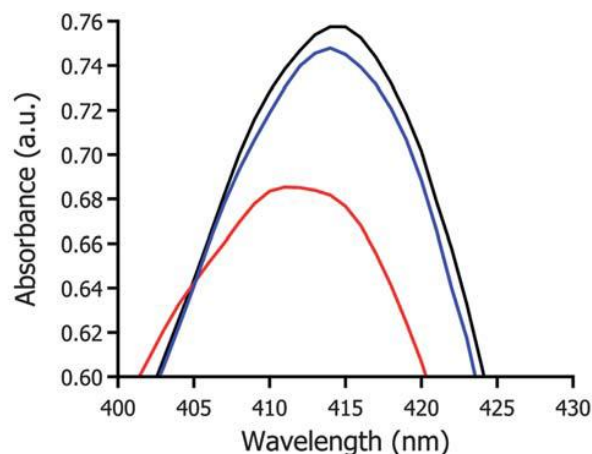


Figure 1.19: DR-UV-vis spectra of solvent-free liquid [C-Mb][S] at 30°C (black), after heating to 155°C (red), and after cooling to 30°C (blue). Samples were incubated for 10 min at each temperature ⁷⁰.

1.4.2.3 Solvent-free liquid protein dynamics

Through dynamic protein modelling and elastic incoherent neutron scattering (EINS), it has been demonstrated that the electrostatically coupled polymer surfactant corona acts as a hydration-shell substitute ^{86,87}. Through molecular dynamic simulations, it was shown that the coronal surfactant shell is able to assemble without disrupting the protein structure and is highly sensitive to the dielectric constant of the surrounding solvent, rearranging into an energy stable conformation when transferred between polar and non-polar solvents, again without protein structure disruption (Figure 1.20A and B). Analysis of atomic mean square displacements (MSD), obtained from EINS of [C-Mb][S] and native myoglobin, it was evident that the motions of the protein within the protein melt were qualitatively similar those of a protein within water. There were some discrepancies in the nearest neighbour centre-centre separations, however, these were thought to be due to a disparity in the weight of the polymer. Whilst the literature value was 737 g.mol⁻¹, used for the molecular dynamics modelling, mass spectrum analysis of the polymer showed a broad spectrum of molecular weights, averaging at 856 g.mol⁻¹. Simulations also displayed a strong adherence of the polymer surfactant to the cationized protein structure, interacting with the protein in a similar fashion to the hydration water.

The molecular dynamic simulations paired with the EINS data lead to the conclusion that the surfactant corona ensures the macromolecular flexibility of the protein hybrid, operating as a plasticizer, affording the dehydrated protein motion and access to different conformational states.

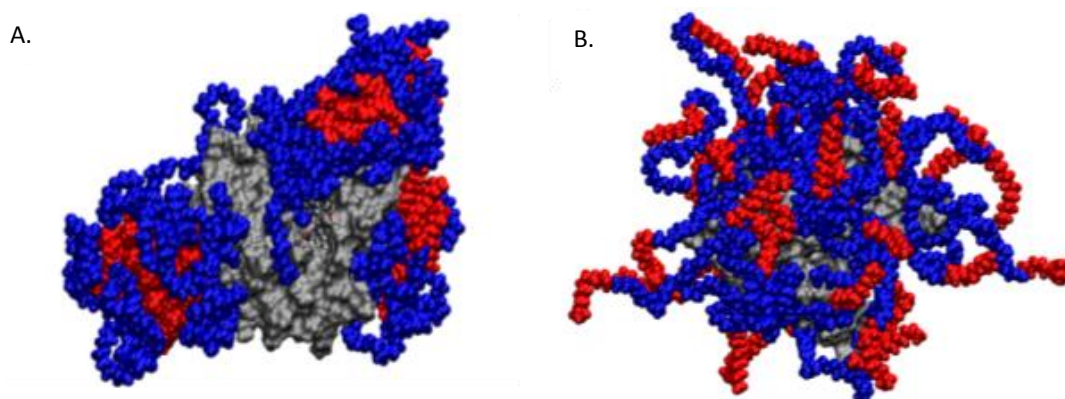


Figure 1.20: Molecular models displaying the arrangement of polymer surfactant molecules after a 100 ns simulation for [C-Mb][S] within **A.** water and **B.** acetonitrile. The **red** areas represent hydrophilic PEG groups and the **blue** areas represent hydrophobic alkyl chains, with the grey areas representing the cationized solvent-accessible sites upon the surface of the protein⁷¹.

1.4.2.4 *Trapping of highly reactive intermediate protein states*

Due to the structural freedom that the surfactant corona of a protein melt affords the protein, acting as a replacement water hydration shell, the solvent-free liquid protein synthesis has been utilised to stabilise intermediate protein folding states that would otherwise aggregate within aqueous media. These intermediate states generally require adverse conditions such as, “extremely low pH or the presence of fluoroalcohols”⁷⁷, to be stabilised. One such example of this is lysozyme, which has a thermally induced intermediate unfolding-state that spontaneously aggregates into amyloid fibres. Through creating a solvent-free liquid lysozyme, [C-Lyz][S], this intermediate state was assessable due to it being trapped in the surfactant corona and therefore unable to aggregate. The intermediate state was discovered to be in equilibrium with the native lysozyme at 78 °C, however over this temperature it transformed into a β -sheet enriched state that did not re-fold once cooled, a characteristic previously demonstrated in other protein melts⁷². This, therefore, displays a 3-stage denaturation pathway which is a shift from the common two-stage denaturation pathway of lysozyme in water (Figure 1.21). The [C-Lyz][S], similar to other protein melts, had a high T_m in contrast with the native lysozyme, displaying the retention of secondary structure at very high temperatures, upon to 210 °C.

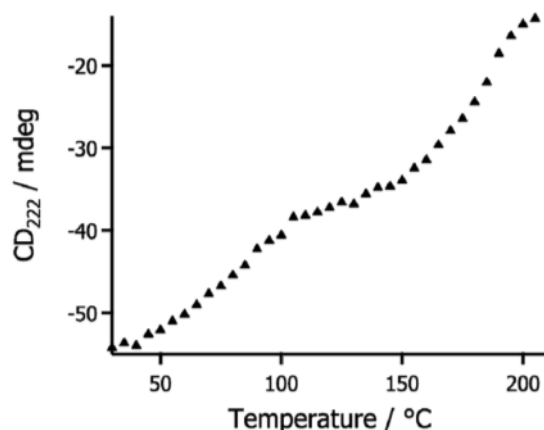


Figure 1.21: Plot of CD peak intensity at 222 nm for solvent-free liquid [C-Lyz][S] at temperatures between 25 and 210 °C at 5 °C intervals showing the presence of a quasi-stable intermediate state in the unfolding pathway between approximately 100 and 150 °C.⁷⁷

1.4.3 Biological function within solvent-free liquid proteins

Given the comparable secondary structure and protein dynamics of the solvent-free liquid proteins to the aqueously solvated native proteins, it would seem logical that the biological function, which is intrinsically linked to the conformation of the protein, would still be present and active. As a proof of concept, the biomolecular functionality of solvent-free liquid myoglobin ([C-Mb][S]) was tested and compared to the functionality of native myoglobin⁷⁰. Firstly deoxy-[C-Mb][S] was successfully formed, with a strong Soret band at 424nm signifying a deoxy Fe(II) porphyrin metallocentre, through the deoxygenation of aqueous met-[C-Mb][S] (followed by the liquid nitrogen freezing and lyophilisation stages of the synthesis). The deoxy-[C-Mb][S] was then subjected to dry environments of; dioxygen, carbon monoxide, and sulfur dioxide, displaying Soret bands of; 418 nm, 417 nm, and Q-bands of 396 nm and 635 nm respectively, with the binding of sulfur dioxide being confirmed through ATR-FTIR. Each of these Soret bands and Q-bands indicate the binding of the gaseous molecule in a conformation similar to that found in the binding of the gaseous molecules within native myoglobin (Figure 1.22). It is to be noted that the time period for the reactions to take place was of the minute time-scale rather the millisecond time-scale as with native Mb. This is due to the high viscosity of the solvent-free liquid protein increasing the diffusion time of the gases and was in line with the estimated diffusion time (17 minutes) for the sample viscosity and thickness. This evidence for biological activity of solvent-free liquid proteins paved the way for more complex solvent-free liquid protein biological functionality to be tested.

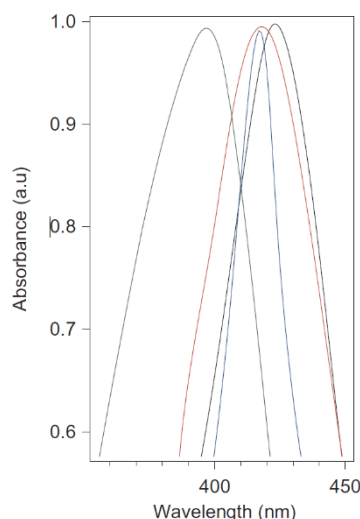


Figure 1.22: Diffuse reflectance UV-Vis spectra showing associated shifts in the Soret band for deoxy-[C-Mb][S₁] (black), oxy-[C-Mb][S₁] (red), CO-[C-Mb][S₁] (blue) and SO₂-[C-Mb][S₁] (green)⁷⁰.

This step forward in the complexity of solvent-free biomolecular functionality was taken through the activity analysis of solvent-free liquid lipase⁶. Lipases catalyse the breakdown of fatty acid esters through a nucleophilic attack upon the substrate which forms an acylated enzyme along with an alcohol by-product. In order to turn over the enzyme, lipases require the nucleophilic action of a water molecule to break down the acylated enzyme, forming a carboxylic acid as a second by-product. The access to the active site of the lipases studied, a mesophilic lipase *rhizomucor miehei* (RML) and a thermophilic lipase *thermomyces languius* (TLL), was controlled by a helical surface-active loop that, when prompted, altered the conformation making the active site solvent accessible. RML and TLL were used to form protein melts, [C-RML][S] and [C-TLL][S], respectively, both displaying native-like structure post synthesis. Through UV-Vis tracking of a liquid phase ester (p-nitrophenyl butyrate; pNPB) that is yellow once hydrolysed, it was evident that aqueous [C-RML][S] and [C-TLL][S] remained active. The cationization stage of the synthesis reduced the activity by 90%, with the surfactant conjugation stage only reducing it by a further 5%. The activity of [C-RML][S] was similar to the activity of aqueous [C-RML][S] being 5% of the native RML, however the activity of [C-TLL][S] was markedly lower than aqueous [C-TLL][S], displaying 0.02% activity when compared to native TLL. The lipase activity, whilst being lower relative to the native species, is also hugely limited in turnover, there being only 20-30 water molecules per protein construct reduces the probably overall turnover number relative to an aqueously solvated lipase.

It should be noted that the extreme decrease in activity of [C-TLL][S] was partially attributed to the geometry of the TLL active-site, being less solvent accessible than RML, making it less probable for the limited number of water molecules to enter the active site. It is also to be noted that the initial enzymatic rate of the lipase melts was somewhat lower than the native and aqueous conjugates, a trait partially attributed to the high viscosity of the lipase melts retarding mass transfer. The evidence of lipase activity in a solvent-free state, albeit limited, provides further evidence for the validity of the protein melt synthesis as a viable surface modification (Figure 1.23A and B).

As there is evidence of native-like secondary structure of protein melts at over 200 °C, the activity of lipase melts was hypothesised to be present at high temperatures. In order to track the activity of [C-RML][S] and [C-TLL][S] melts at higher temperatures, a solid phase ester with a high boiling point that changes colour upon hydrolysis was required (p-nitrophenyl palmitate; pNPPal, was used as the boiling point of pNPB is 125 °C, changing from clear to yellow). The UV-Vis tracking assays were conducted from 30-120 °C and displayed a relatively constant increase in enzymatic rate as the temperature increased, 88-fold and 45-fold for [C-RML][S] and [C-TLL][S] respectively (Figure 1.23C). This increase in enzymatic rate was attributed to a higher kT along with a reduction in viscosity of the protein melt due to the temperature increase. The lower viscosity of the higher temperature melt having a more rapid mass transfer than the lower temperature highly viscous melt.

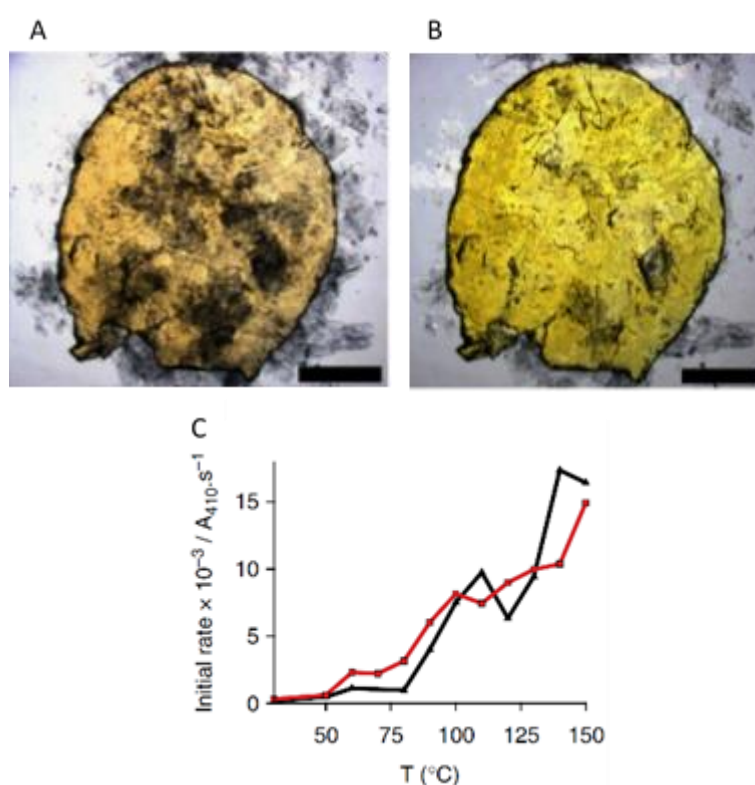


Figure 1.23: Microscopy images displaying [C-RML][S] and desiccated pNPB substrate **A.** immediately after contact and **B.** after incubation for 24 hr at 50°C. **C.** Plot showing a positive correlation between an increase in the initial reaction rate with an increase in temperature for the reaction between pNPPal and solvent-free liquid [C-RML][S]_i (black) or solvent-free liquid [C-TLL][S]_i (red), taken at 10 °C intervals⁶.

1.4.4 Solvent-free liquid proteins solvated in ionic liquids

Ionic liquids are increasingly being looked at as a viable alternative to organic solvents, being chemically and physically 'tunable', widening their scope of use within industrial processes. However, dry proteins have a poor solubility within ionic solvents, with native proteins also being thermally limited to low temperature environments, restricting their use by lowering the potential reaction kinetics. Recently, a combination of the highly thermally stable solvent-free liquid proteins and the highly thermally stable tunable ionic liquids was investigated by Brogan *et al*, with myoglobin⁸⁵ and

glucosidase^{88,89} being the proteins studied. The surfactant corona surrounding the protein was shown to increase the solubility of the protein within the ionic liquid, also increasing the homogeneity of protein mixing. The protein/ionic-liquid solution was shown to not only retain the high level of native secondary structure seen in the solvent-free liquid myoglobin but to increase this level towards a more native-like structure. Being solvated within the ionic liquid also preserved the high thermal stability of the modified proteins, increasing the half denaturation (a measure of stability) by 55 °C and 66 °C relative to the native myoglobin and glucosidase (respectively) in aqueous conditions. This increase took the half denaturation past the boiling point of water, on indication of how beneficial the anhydrous ionic solvents are. Glucosidase is an enzyme that catalyses the hydrolysis of oligosaccharides within a cellulase enzyme cascade that breaks down cellulose into sugar. This a set of reactions that is an especially useful due to cellulose being an abundant natural and renewable material along with being within many waste products⁹⁰. Cellulose can be pre-treated within ionic liquids to reduce the crystallinity and making it more susceptible to hydrolysis through the cellulase cascade. Incorporating solvent-free liquid glucosidase into an ionic liquid displayed a retention of aqueous buffered native-like levels of activity, however, increasing the temperature up to 110 °C increased the enzymatic activity by almost 30-fold, a marker improvement relative to the lower temperature (Figure 1.24). Being soluble within ionic liquids, with retention of a high-level of native-like structure and thermal resistance, along with an increase in activity, paves the way for enzymes dissolved within ionic liquids to be incorporated into a greater number of processes and reactions, potentially increasing the output.

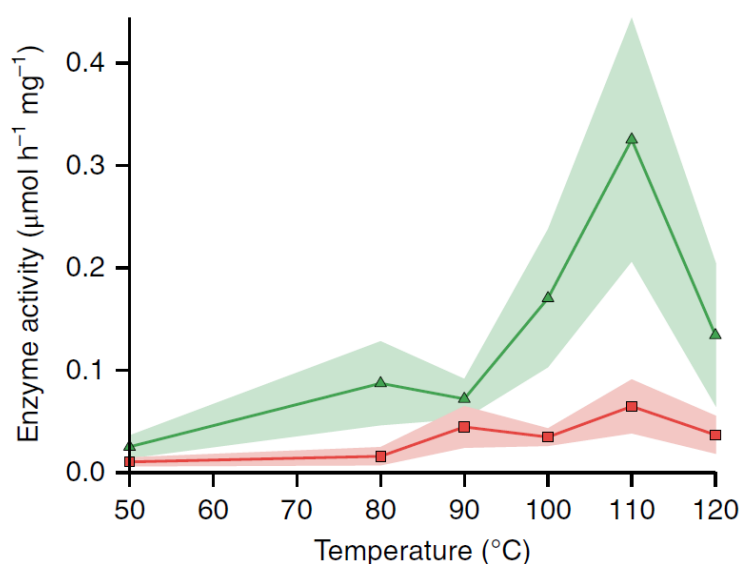


Figure 1.24: Plot of enzyme activity against temperature for solvent-free liquid glucosidase in the ionic liquid [emim][EtSO₄] (emim = 1-ethyl-3-methylimidazolium and EtSO₄ = ethyl sulfate), where cellobiose (green) and cellulose were the substrates (red). The shaded areas of the graph represent the standard deviations ($n \geq 6$)⁹⁰.

1.5 Introduction to proteases

Proteases, proteinases, or peptidases are a protein super-family with the common function of performing proteolysis. Proteolysis is the catalysis of the hydrolysis reaction of the peptide bond, which links two amino acids together, and is therefore a mechanism for either the partial or complete degradation of a protein polypeptide chain^{91, 92}. In addition to catalysing the breakdown of peptide bonds, if working within the correct environmental conditions, proteases can also function as peptide synthesisers, catalysing the formation of peptide bonds^{93, 94}.

Proteases are found within every living organism, both prokaryotes and eukaryotes, playing an integral role on a wide variety of levels: from within the physiology of the cell, aiding in many cellular functions such as apoptosis⁹⁵; to the organism level, producing regulatory and inflammatory cascades⁹⁶. As proteases are intrinsically linked to many aspects of the life cycle of many organisms on the cellular level, they are associated with a large number of processes or degradation of processes that give rise to diseases. They have therefore become a target for development with potential use as a drug treatment aimed at combating diseases such as type 2 diabetes, multiple myeloma, HIV and hepatitis C virus infections⁹⁷. As well as inhibiting proteases within the body to cure illnesses, proteases used in growth mechanisms within bacteria can be targeted, stifling or preventing growth of bacterial colonies. Due to the highly selective cleavage sites and the potential to modify the surface and/or active site, proteases could be employed as an alternative to antibiotics or used to circumvent the current antibiotic resistant bacteria issues. The protease could target specific structures within the bacteria, such as a particular DNA sequence, and cleave them, effectively destroying the bacteria^{97, 98, 99}.

In addition to being physiologically necessary, proteases are also widely utilised within industry, increasing the rate of a multitude of reactions. For example: chymosin, the primary active ingredient in rennet, and cleaves a very specific bond within κ -casein which has the effect of curdling milk¹⁰⁰. Other well-known examples are the collection of proteases that are known to age and tenderise red meat¹⁰¹ and the use of papaya proteinase for cold proofing beer¹⁰². There are many proteases that perform a variety of functions due to the potential highly selective nature of bond cleavage. The protease mechanism has evolved convergently through over 60 separate evolutionary lines¹⁰³. As so many different proteases exist it is necessary to logically group them to aid in their description and practical use. Many different classification methods are used to name, describe, and then group proteases including structure, catalytic residue and substrate specificity.

1.5.1 General protease mechanisms

The protease family is formed of six major groups named after the active catalytic residue/inorganic molecule within the protein¹⁰⁴: serine, threonine, cysteine, aspartate, and glutamic acid, along with

metalloproteases which utilise a metal within the active site. Each of these active residues/ions form a pivotal role in the break of the scissile bond, the bond being broken by the enzyme¹⁰⁵. These six groups can be split into two divisions; aspartic, glutamic, and metalloproteases; and serine, cysteine, and threonine. The two divisions are differentiated *via* the mechanism used to catalyse the hydrolysis of the scissile bond; either through the active residue activating a water molecule which in turn performs a nucleophilic attack upon the scissile peptide bond (Figure 1.25) or through the active residue within a catalytic triad performing the nucleophilic attack upon the scissile peptide bond (Figure 1.26). The latter mechanism involves breaking the peptide bond and forming an acyl-enzyme intermediate, with the active residue covalently bonding to the intermediate. This is then followed by the hydrolysis of this bond *via* nucleophilic attack of an activated water molecule, releasing the product.

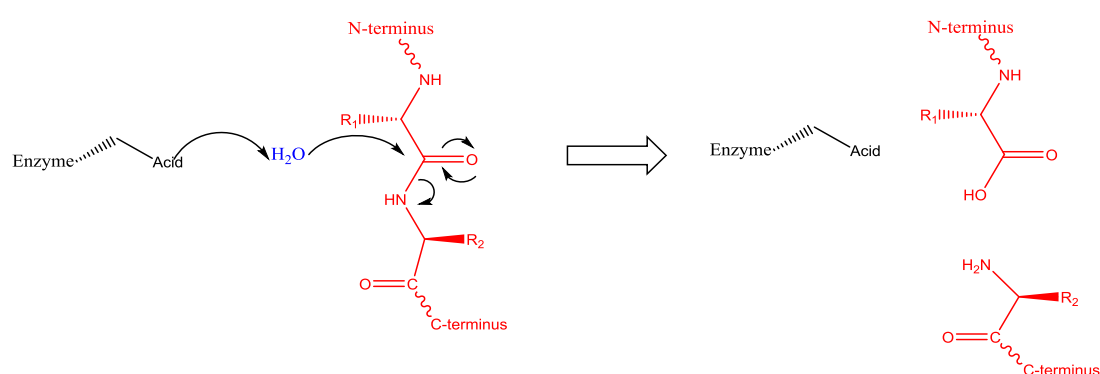


Figure 1.25: A schematic displaying the general mechanism of an activated water-based protease (black) cleaving a general polypeptide substrate (red).

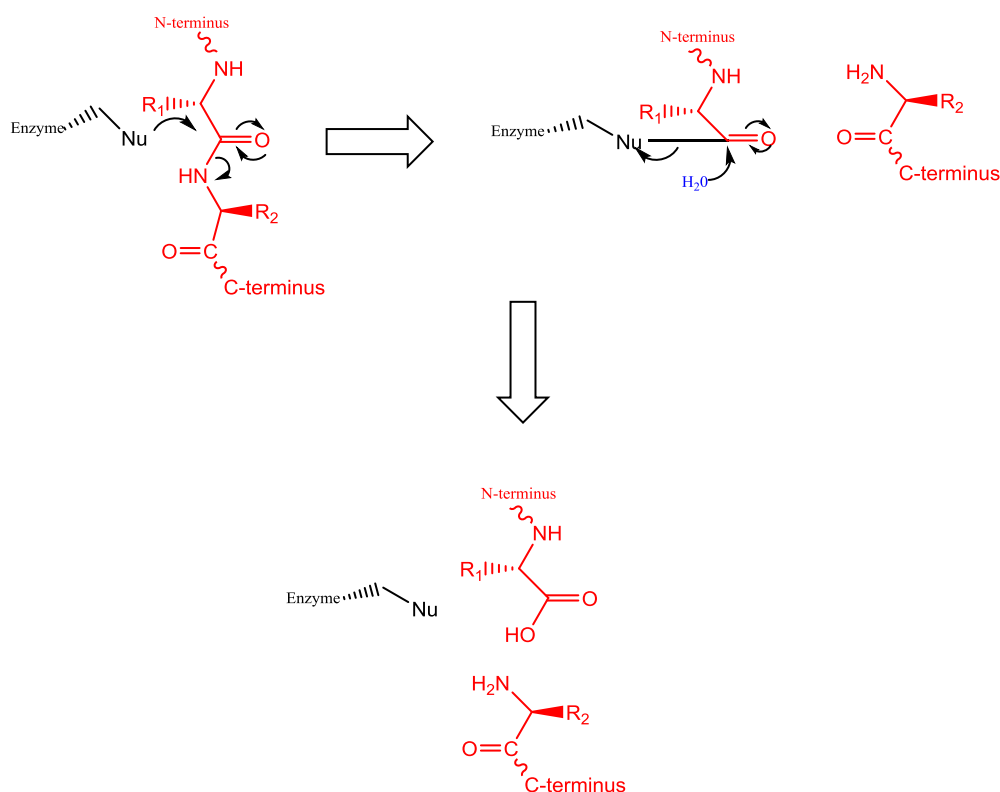


Figure 1.26: A schematic displaying the general mechanism of a nucleophilic residue-based protease (black) cleaving a general polypeptide substrate (red).

1.5.2 The protease active site

Each of the six major protease groups can be further divided into sub-groups, with their nomenclature and mode of action based around the main catalytic residue present¹⁰⁴. Each sub-group is distinguished by the specificity of the cleavage site, *i.e.* where the protease can cleave along a polypeptide chain, whilst retaining the same active catalytic residue. The specificity of each protease sub-group derives from two major areas; the residues within the active site known as the catalytic triad or diad; and the residues located around the active site, an area known as the substrate binding cleft, which form hydrogen bonds with specific sequences of substrate residues.

1.5.2.1 Catalytic site

Each protease has an active site which is the location of the main catalytic residues utilised for the hydrolysis of polypeptides, often formed of catalytic dyads or triads¹⁰⁶. A catalytic dyad or triad works through the inter-residue interactions of two or three amino acid residues, respectively, that are in a close, specific conformation. These residues interact with the substrate and molecules in the surrounding area to aid in the creation and stabilisation of intermediate products and the eventual creation of the final products. It is common in many proteases to form acid-base-nucleophile dyads and triads^{107, 108}. In most catalytic dyads and triads there is one primary residue that interacts directly with the substrate or with the attacking species, and then one or two secondary residues in close

proximity which interact with the primary residue in such a way that creates a more reactive species, along with a more stable intermediate product.

1.5.2.2 *Binding cleft*

The binding cleft is formed from at least one residue but is often formed from more, with proteases incorporating a higher number of residues on the binding cleft having a more specific cleavage, such as tobacco etch virus (TEV) protease which requires a specific 7 amino acid residue sequence¹⁰⁹. The residues present on the binding cleft can also be utilised within the catalytic process, aiding in the stabilisation of the substrate during hydrolysis (Figure 1.10B). However, it is not only residues on the binding cleft that can aid stabilisation, residues present on the peptide backbone, specifically residues proximal to the active site, are able to aid in stabilisation¹¹⁰. These residues can form O-H hydrogen bonds with the scissile carbonyl oxygen, known as the oxyanion hole or oxyanion binding site. Not only does this stabilise the substrate in a tetrahedral form, but it also polarises the carbonyl, moving negative charge to the oxygen¹¹¹. This activates the carbonyl carbon, making it more susceptible to nucleophilic attack during the catalytic process. The additional hydrogen bonds formed between the peptide backbone and the protein being hydrolysed also allows for greater stabilisation of any intermediate species that form. These are thought to be a major contributor to the increased rate of reaction when using proteases. The oxyanion hole is not found in all members of the protease superfamily but is almost ubiquitous in the serine protease clan.

Cleavage specificity can be due to a number of different reasons, and often a number of these reasons are combined for each sub-group; residue type, C/N terminal, size and charge of side chains, size and charge of neighbouring residues, the residue sequence either side of the scissile bond. It can also be due to whether the protease is cleaving a terminal group (exoprotease) or whether the cleavage site is within a polypeptide chain (endoprotease). Protein residue recognition is therefore a key, and is also the initial, step in the proteolysis mechanism.

1.5.3 Specific protease mechanisms

As previously stated, there are two overarching protease mechanisms; one that involves the primary catalytic residue interacting directly with the polypeptide chain; and a second that involves the primary catalytic residue interacting with an attacking species. These two mechanisms are used to generalise the protease substrate interactions as they are a good outline as to how a protease can interact with a substrate.

The first catalytic mechanism begins with a nucleophilic attack on the scissile carbonyl carbon of the substrate, the attack coming from a water molecule that is activated through interactions with the primary residue. There are two major groups associated with this mechanism; aspartate proteases and

metalloproteases, both utilising a residue within the catalytic site to de-protonate a water molecule¹¹². This water molecule then initiates hydrolysis *via* a nucleophilic attack on the scissile carbonyl carbon on the substrate. After this initial activation stage, the mechanisms of each major group then deviate, as detailed below, both forming two products; N- and C-terminus peptides.

1.5.3.1 Aspartate and metallo- protease mechanism

Aspartate proteases are members of a group of proteases that utilise aspartate residues to catalyse the hydrolysis of protein polypeptide bonds. They do this through the use of a catalytic dyad active site, the dyad being two aspartate residues in close proximity, to de-protonate a water molecule that is held within the active site of the protease, as seen in Figure 1.27.

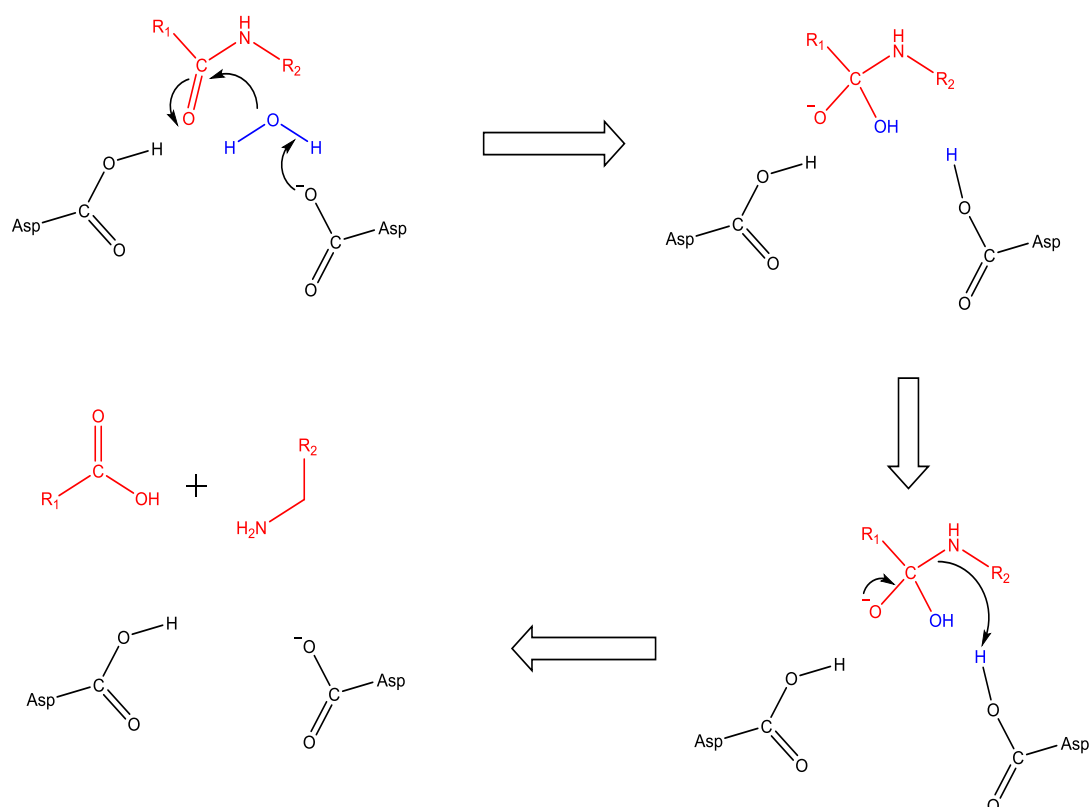


Figure 1.27: A schematic displaying the mechanism for an aspartate protease (black) cleaving a general polypeptide substrate (red) with the aid of a water molecule (blue).

Metalloproteases are, like aspartate proteases, members of a clan that hydrolyse protein polypeptide bonds, however rather than aspartate residues, metalloproteases utilise a metal ion, along with a number of amino acid residues, which are located within the catalytic site of the protease¹¹³. The metal ion, commonly a zinc ion (Zn^{2+}), is bound in an almost tetrahedral formation, associating with three nearby charged residues and a local water molecule.

1.5.3.2 Serine protease mechanism

The second mechanism, often termed serine-like, involves the breaking of the scissile bond *via* nucleophilic attack through the primary residue followed by the formation of an intermediate product¹¹⁴. The primary residue initially forms a covalent bond with the polypeptide chain creating a stable tetrahedral intermediate species. A suitable leaving group, being the N-terminus product, is created through interactions with surrounding residues and leaves, initiating the breakdown of the tetrahedral intermediate and the creation of an acyl-enzyme intermediate. A second nucleophilic attack is required to breakdown this intermediate species further, leaving the two products and the original primary catalytic residue. This attack takes the form of a deprotonated water molecule, activated *via* an interaction with a nearby residue, creating a second tetrahedral intermediate. Through additional interactions with nearby residues, a leaving group, being the C-terminus product, is created and released. An example of a serine-like catalytic triad, e.g. serine catalytic triad of serine, histidine, and aspartic acid residues, is detailed in (Figure 1.28), with a general polypeptide as the substrate and with the aspartic acid residue is omitted for clarity¹¹⁵.

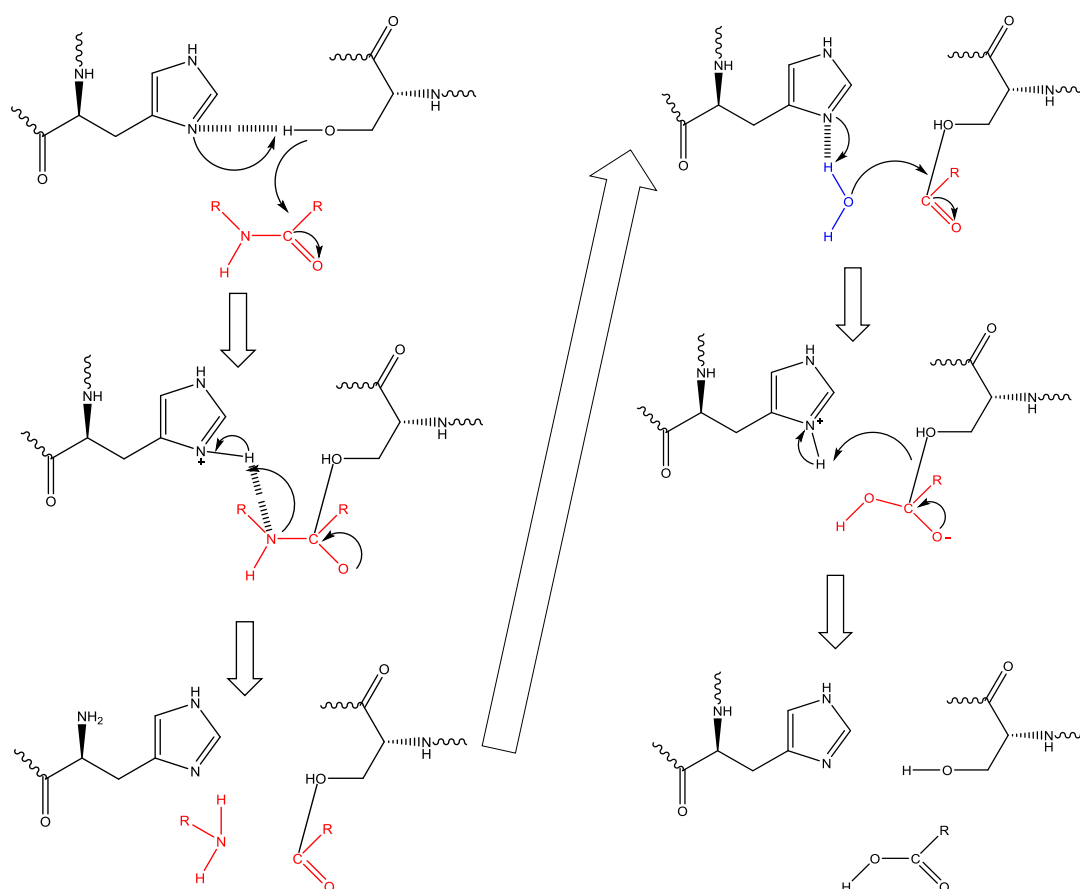


Figure 1.28: A schematic displaying the mechanism for a Serine/Histidine/Aspartate catalytic triad protease (black) cleaving a general polypeptide substrate (red) with the aid of a water molecule (blue).

The initial steps of the catalytic process (not displayed in *figure 19*) are the binding of the substrate to the protease's binding cleft and the activation of the carbonyl carbon scissile peptide bond *via* the oxyanion hole. Omitted from the diagram is the aspartic acid residue. The aspartic acid residue is located near the histidine residue and aids in the orientation of this residue, and is also in a proximity that allows it to neutralize the charge build up that develops on the histidine residue during the catalytic process¹¹⁶.

To initiate the hydrolysis the serine residue performs a nucleophilic attack upon the activated carbonyl carbon on the polypeptide chain forming a stable tetrahedral intermediate product. However, serine is known to be a poor nucleophile. For the serine to increase in nucleophilicity, the histidine, correctly orientated by the aspartate residue, acts as a general base, deprotonating the alcohol on the serine residue, increasing its nucleophilic strength. The deprotonated serine can now initiate the nucleophilic attack upon the substrate forming a tetrahedral intermediate product. There is no obvious leaving group for this stable intermediate product, but the now protonated histidine, an imidazolium ion, can act as a general acid, accepting a lone pair of electrons from the present amino group. The imidazolium ion therefore protonates the amino group, creating a leaving group which initiates the breakdown of the tetrahedral intermediate to an ester-enzyme intermediate and produces the N-terminus product.

To breakdown the ester intermediate into the C-terminal product and free the enzyme, a second nucleophilic attack is required. The nucleophilic species takes the form of a de-protonated water molecule and attacks the still active carbonyl carbon forming a second stable tetrahedral intermediate. The histidine side chain, which has now returned to its original form, is able to act, again, as a general base, de-protonating the water molecule forming a second imidazolium ion. Once the tetrahedral intermediate has formed, the imidazolium ion, acting as a general acid, protonates the serine-oxygen creating a good leaving group in the form of the C-terminus product.

Serine-like proteases have been split into 16 families which can be further divided into sub-categories, however, the major groups are trypsin-like, chymotrypsin-like, elastase-like and subtilisin-like^{[93] [94]}. All of the aforementioned groups have a Ser/His/Asp catalytic triad however, other serine proteases may have variations within the triads or even have dyads. Each group has a varied tertiary structure and therefore binding cleft, along with varying sized catalytic sites, meaning they all have different substrate specificities.

It is through this serine-like catalytic mechanism that the proposed protease utilises in order to hydrolyse polypeptide bonds. The active site, as detailed above, contains a ser/his/asp triad with serine being the primary residue.

1.5.4 Proteinase K

Proteinase K (PK) from *Engyodontium album* (formerly *Tritirachium album*) is a subtilisin-like serine endoprotease containing 277 amino acids (Mr 28,930 Da)¹¹⁷ that preferentially cleaves the bond proximate to carboxyl groups of hydrophobic residues (aliphatic and aromatics), although it is known to be a broad-spectrum protease^{118, 119}. There are four salt bridges present in PK structure (Asp 187 – Arg 12, Asp 117 – Arg 121, Asp 98 – Lys 94, and Asp 184 – Arg 188)¹²⁰ that contribute to the stabilisation of the tertiary protein structure that displays an extended central parallel β -sheet that is decorated by six α -helices, three short antiparallel β -sheets, and 18 β -turns¹²¹ (Figure 1.29). The catalytic triad involves Ser 224, His 69, and Asp 39, forming a complex hydrogen bonding network along with Asn 161 and the water-containing oxyanion hole located within the active site (Figure 1.30). These water molecules are known to be significant in the conformational integrity of the active site, connecting areas through hydrogen bonding¹²¹. There are two calcium (Ca^{2+}) site located near the active site that are required for the correct folding and stabilisation of the polypeptide chain¹²². However, they are not required for catalytic activity of the protease, with removal leaving the proteolytic activity unchanged¹²². Proteinase K is stable and active over a wide range of pH values (4-12) and temperatures but is optimally active at pH 8 and is known to retain activity in the presence of surfactants such as SDS^{123, 124}.

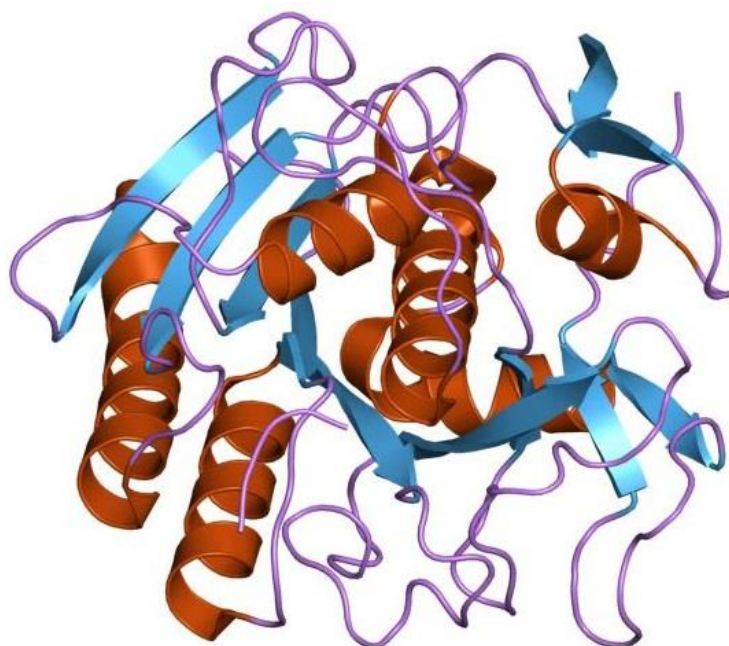


Figure 1.29: Ribbon representation of proteinase K, with α -helices (orange), β -sheets (blue), and other contents (purple)¹²⁵.

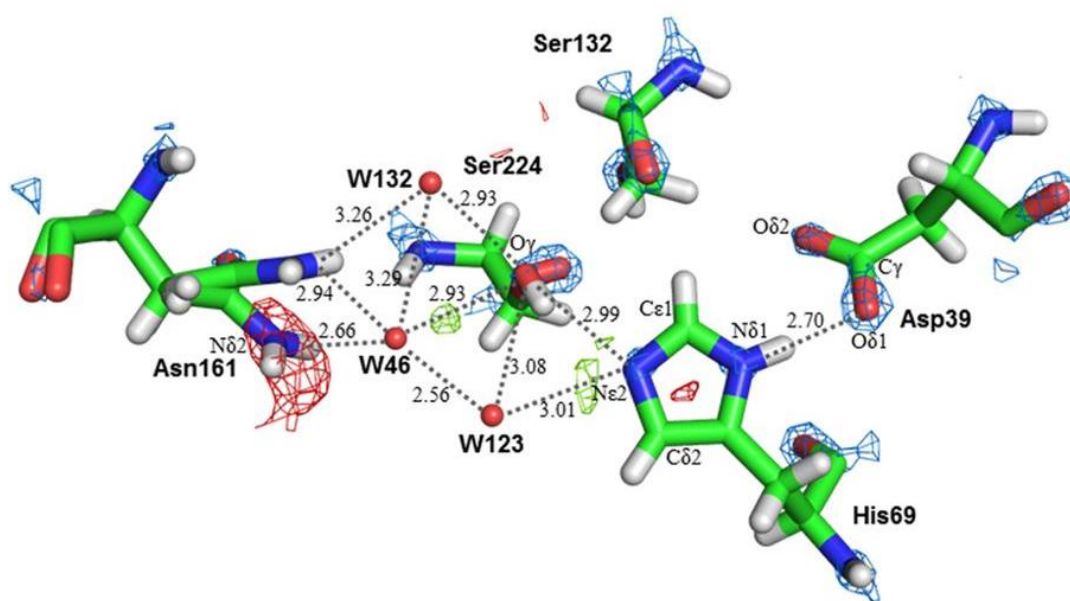


Figure 1.30: The active site of proteinase K as determined via X-ray crystallography, displaying the hydrogen bonds (dotted lines) and the lengths of those hydrogen bonds (Å) between each residue as well as the water molecules within the active site (red balls). The σ_A -weighted 2mFo-DFc maps contoured at 4.5 σ are shown in blue meshes. The σ_A -weighted mFo-DFc maps contoured at 2.5 σ and -2.5 σ are shown in green and red meshes, respectively¹²⁶.

1.6 References

1. Boutureira, O. & Bernardes, G. J. L. Advances in chemical protein modification. *Chem. Rev.* **115**, 2174–95 (2015).
2. Rozbeský, D. *et al.* Impact of Chemical Cross-Linking on Protein Structure and Function. *Anal. Chem.* **90**, 1104–1113 (2018).
3. Perriman, A. W., Cölfen, H., Hughes, R. W., Barrie, C. L. & Mann, S. Solvent-free protein liquids and liquid crystals. *Angew. Chemie - Int. Ed.* **48**, 6242–6246 (2009).
4. Perriman, A. W. *et al.* Reversible dioxygen binding in solvent-free liquid myoglobin. *Nat. Chem.* **2**, 622–6 (2010).
5. Chemistry, O. M. Support: Isolation of a highly reactive β -sheet rich intermediate of lysozyme in a solvent-free liquid phase. Alex P. S. Brogan, Kamendra P. Sharma, Adam W. Perriman * and Stephen Mann. **2**, 1–11 (2013).
6. Brogan, A. P. S., Sharma, K. P., Perriman, A. W. & Mann, S. Enzyme activity in liquid lipase melts as a step towards solvent-free biology at 150 °C. *Nat. Commun.* **5**, 6058 (2014).
7. Harris, L. & Birch, T. Zwitterions: Proof of the zwitterion constitution of the amino-acid molecule. II. Amino-acids, polypeptides, etc., and proteins as zwitterions, with instances of non-zwitterion ampholytes. *Biochem. J.* **24**, 1080–97 (1930).
8. Esposito, L., De Simone, A., Zagari, A. & Vitagliano, L. Correlation between ω and ψ dihedral angles in protein structures. *J. Mol. Biol.* **347**, 483–487 (2005).
9. Ramakrishnan, C. Stereochemical criteria for polypeptide and protein chain conformations - Part I. Evaluation of helical parameters. *Proc. Indian Acad. Sci. - Sect. A* **59**, 327–343 (1964).
10. Hollingsworth, S. A. & Karplus, P. A. A fresh look at the Ramachandran plot and the occurrence of standard structures in proteins. *Biomol. Concepts* **1**, 271–283 (2010).
11. Kabsch, W. & Sander, C. Dictionary of protein secondary structure: pattern recognition of hydrogen bonded and geometrical features. *Biopolymers* **22**, 2577–2637 (1983).
12. Berg JM, Tymoczko JL, S. L. *Biochemistry*. (W H Freeman, 2002).
13. Harrison, S. C. & Blout, E. R. Reversible Conformational Changes and Apomyoglobin of Myoglobin. **240**, (1965).
14. Fedotova, M. V. & Dmitrieva, O. A. Characterization of selective binding of biologically relevant inorganic ions with the proline zwitterion by 3D-RISM theory. *New J. Chem.* **39**, 8594–8601

- (2015).
15. Gutteridge, A. & Thornton, J. Conformational change in substrate binding, catalysis and product release: An open and shut case? *FEBS Lett.* **567**, 67–73 (2004).
 16. A.C, H. & B.W, M. Thermolysin in the absence of substrate has an open conformation. *Acta Crystallogr.* **58**, 1002–1007 (2002).
 17. Mihailescu, M. R. & Russu, I. M. A signature of the T → R transition in human hemoglobin. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 3773–7 (2001).
 18. Paoli, M., Liddington, R., Tame, J., Wilkinson, A., Dodson, G. Crystal Structure of T State Haemoglobin with Oxygen Bound at All Four Haems. *J.Mol.Bio* **256**, 775 (1996).
 19. Pace, C. N., Treviño, S., Prabhakaran, E. & Scholtz, J. M. Protein structure, stability and solubility in water and other solvents. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **359**, 1225–34; discussion 1234–5 (2004).
 20. Pace CN. The stability of globular proteins. *CRC Crit. Rev. Biochem.* **3**, 1–43 (1975).
 21. Breiten, B. *et al.* Water networks contribute to enthalpy/entropy compensation in protein-ligand binding. *J. Am. Chem. Soc.* **135**, 15579–15584 (2013).
 22. Charton, M. & Charton, B. I. The structural dependence of amino acid hydrophobicity parameters. *J. Theor. Biol.* **99**, 629–644 (1982).
 23. Chen, J. & Stites, W. E. Packing is a key selection factor in the evolution of protein hydrophobic cores. *Biochemistry* **40**, 15280–15289 (2001).
 24. Wedemeyer, W. J., Welker, E., Narayan, M. & Scheraga, H. A. Disulfide bonds and protein folding. *Biochemistry* **39**, 4207–4216 (2000).
 25. Walker, L. C. & Levine, H. The cerebral proteopathies. *Neurobiol. Aging* **21**, 559–561 (2000).
 26. Bobadilla, J. L., Macek, M., Fine, J. P. & Farrell, P. M. Cystic fibrosis: A worldwide analysis of CFTR mutations - Correlation with incidence data and application to screening. *Hum. Mutat.* **19**, 575–606 (2002).
 27. Stephanopoulos, N. & Francis, M. B. Choosing an effective protein bioconjugation strategy. *Nat. Chem. Biol.* **7**, 876–84 (2011).
 28. Prabakaran, S., Lippens, G., Steen, H. & Gunawardena, J. Post-translational modification: Nature's escape from genetic imprisonment and the basis for dynamic information encoding. *Wiley Interdiscip. Rev. Syst. Biol. Med.* **4**, 565–583 (2012).
 29. Burnett, G. & Kennedy, E. P. The enzymatic phosphorylation of proteins. *Cancer* 969–980

- (1954).
30. Lieza M. Danan, Zhihao Yu, Peter J. Ludden, Weitao Jia, Kevin L. Moore, and J. A. L. Catalytic Mechanism of Golgi-resident Human Tyrosylprotein Sulfotransferase-2: A Mass Spectrometry Approach. **21**, 1633–1642 (2010).
 31. Nadolski, M. J. & Linder, M. E. Protein lipidation. *FEBS J.* **274**, 5202–10 (2007).
 32. Resh, M. D. Fatty acylation of proteins: new insights into membrane targeting of myristoylated and palmitoylated proteins. *Biochim. Biophys. Acta - Mol. Cell Res.* **1451**, 1–16 (1999).
 33. Casey, P. Protein lipidation in cell signaling. *Science (80-)*. **268**, 221–225 (1995).
 34. Booth, P. J. The trials and tribulations of membrane protein folding in vitro. *Biochim. Biophys. Acta - Biomembr.* **1610**, 51–56 (2003).
 35. Pufall, M. A. & Graves, B. J. Autoinhibitory Domains: Modular Effectors of Cellular Regulation. *Annu. Rev. Cell Dev. Biol.* **18**, 421–462 (2002).
 36. Sung, W. S., Park, Y., Choi, C. H., Hahm, K. S. & Lee, D. G. Mode of antibacterial action of a signal peptide, Pep27 from *Streptococcus pneumoniae*. *Biochem. Biophys. Res. Commun.* **363**, 806–810 (2007).
 37. Pauling, L., Corey, R. B. & Branson, H. R. The structure of proteins: Two hydrogen-bonded helical configurations of the polypeptide chain. *Proc. Natl. Acad. Sci.* **37**, 205–211 (1951).
 38. Kemmler, W., Peterson, J. D. & Steiner, D. F. On the Conversion of Proinsulin to Insulin. *J. Biol. Chem.* **246**, 6786–6791 (1971).
 39. Jutras, I. & Reudelhuber, T. L. Prorenin processing by cathepsin B in vitro and in transfected cells. *FEBS Lett.* **443**, 48–52 (1999).
 40. Oláh, A. & Romics, L. Enteral nutrition in acute pancreatitis: A review of the current evidence. *World J. Gastroenterol.* **20**, 16123–16131 (2014).
 41. Kalia, J. and R. T. R. Advances in Bioconjugation. *Curr Org Chem* **14**, 138–147 (2010).
 42. Hutchison, C. A. *et al.* Mutagenesis at a specific position in a DNA sequence. *J. Biol. Chem.* **253**, 6551–6560 (1978).
 43. Fan, L. T., Lee, Y.-H. & Beardmore, D. H. Mechanism of the enzymatic hydrolysis of cellulose: Effects of major structural features of cellulose on enzymatic hydrolysis. *Biotechnol. Bioeng.* **22**, 177–199 (1980).
 44. al, S. *et.* D-peptide Protease Resistance. at <https://sites.tufts.edu/dpeptide172/mechanism/>
 45. Desantis, G. & Jones, J. B. Chemical modification of enzymes for enhanced functionality. *Curr.*

-
- Opin. Biotechnol.* **10**, 324–330 (1999).
46. Einstein, a. Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Ann. d. Phys.* **322**, 549–560 (1905).
 47. Iyer, P. V. & Ananthanarayan, L. Enzyme stability and stabilization—Aqueous and non-aqueous environment. *Process Biochem.* **43**, 1019–1032 (2008).
 48. Zaks, a & Klivanov, a M. Enzymatic catalysis in nonaqueous solvents. *J. Biol. Chem.* **263**, 3194–201 (1988).
 49. Kumar, S., Tsai, C.-J. & Nussinov, R. Factors enhancing protein thermostability. *Protein Eng. Des. Sel.* **13**, 179–191 (2000).
 50. Sawle, L. & Ghosh, K. How do thermophilic proteins and proteomes withstand high temperature? *Biophys. J.* **101**, 217–227 (2011).
 51. Chan, C.-H., Wilbanks, C. C., Makhatadze, G. I. & Wong, K.-B. Electrostatic contribution of surface charge residues to the stability of a thermophilic protein: benchmarking experimental and predicted pKa values. *PLoS One* **7**, e30296 (2012).
 52. Kluger, R. & Alagic, A. Chemical cross-linking and protein-protein interactions-a review with illustrative protocols. *Bioorg. Chem.* **32**, 451–72 (2004).
 53. López-Gallego, F. *et al.* Enzyme stabilization by glutaraldehyde crosslinking of adsorbed proteins on aminated supports. *J. Biotechnol.* **119**, 70–75 (2005).
 54. Littlechild, J. A. Enzymes from Extreme Environments and Their Industrial Applications. *Front. Bioeng. Biotechnol.* **3**, 1–9 (2015).
 55. Sivapragasam, M., Moniruzzaman, M. & Goto, M. Recent advances in exploiting ionic liquids for biomolecules: Solubility, stability and applications. *Biotechnol. J.* **11**, 1000–1013 (2016).
 56. Schröder, C. Proteins in Ionic Liquids: Current Status of Experiments and Simulations. *Top. Curr. Chem.* **375**, (2017).
 57. Baker, G. a, Baker, S. N., Pandey, S. & Bright, F. V. An analytical view of ionic liquids. *Analyst* **130**, 800–8 (2005).
 58. Lui, M. Y. *et al.* Salts dissolved in salts: Ionic liquid mixtures. *Chem. Sci.* **2**, 1491–1496 (2011).
 59. Saha, D. & Mukherjee, A. Effect of water and ionic liquids on biomolecules. 795–808 (2018). doi:10.1007/s12551-018-0399-2
 60. Kumar, A. & Venkatesu, P. Overview of the stability of α -chymotrypsin in different solvent media. *Chem. Rev.* **112**, 4283–307 (2012).

61. Schein, C. H. Solubility as a Function of Protein Structure and Solvent Components. *Bio/Technology* **8**, 308 (1990).
62. Wang, Y. & Caruso, F. Mesoporous silica spheres as supports for enzyme immobilization and encapsulation. *Chem. Mater.* **17**, 953–961 (2005).
63. Tosa T, Mori T, Fuse N, C. I. Studies on continuous enzyme reactions. I. Screening of carriers for preparation of water-insoluble aminoacylase. *Enzymologia* **31**, 214–224 (1966).
64. Takatsuji, Y. *et al.* Solid-support immobilization of a ‘swing’ fusion protein for enhanced glucose oxidase catalytic activity. *Colloids Surfaces B Biointerfaces* **112**, 186–191 (2013).
65. Bickerstaff, G. Immobilization of Enzymes and Cells. **1**, (1996).
66. Heck, T., Faccio, G., Richter, M. & Thöny-Meyer, L. Enzyme-catalyzed protein crosslinking. *Appl. Microbiol. Biotechnol.* **97**, 461–475 (2013).
67. Jevsevar, S., Kunstelj, M. & Porekar, V. G. PEGylation of therapeutic proteins. *Biotechnol. J.* **5**, 113–28 (2010).
68. Veronese, F. M. & Pasut, G. PEGylation , successful approach to drug delivery REVIEWS. **10**, 1451–1458 (2005).
69. Rodríguez-Martínez, J. a, Rivera-Rivera, I., Solá, R. J. & Griebenow, K. Enzymatic activity and thermal stability of PEG- α -chymotrypsin conjugates. *Biotechnol. Lett.* **31**, 883–7 (2009).
70. Perriman, A. W. *et al.* Reversible dioxygen binding in solvent-free liquid myoglobin. *Nat. Chem.* **2**, 622–6 (2010).
71. Nanoscience, B., Perriman, A. W. & Mann, S. Liquid Proteins — A New Frontier for Biomolecule-Based Nanoscience. **5**, 6085–6091 (2011).
72. Brogan, A. P. S., Siligardi, G., Hussain, R., Perriman, A. W. & Mann, S. Hyper-thermal stability and unprecedented re-folding of solvent-free liquid myoglobin. *Chem. Sci.* **3**, 1839 (2012).
73. Hagen, M. H. J., Meijer, E. J. & Moo, G. C. A. M. Does C60 have a liquid phase ? Enhanced reactivity of fullerene cations containing adjacent pentagons. **365**, 0–1 (1993).
74. Pj, D. & Sc, W. Protein modification by thermal processing. **53**, 102–105 (1998).
75. Rodriguez, R., Herrera, R., Archer, L. a. & Giannelis, E. P. Nanoscale Ionic Materials. *Adv. Mater.* **20**, 4353–4358 (2008).
76. Scott, E. E., Gibson, Q. H. & Olson, J. S. Mapping the pathways for O₂ entry into and exit from myoglobin. *J. Biol. Chem.* **276**, 5177–88 (2001).
77. Brogan, A. P. S., Sharma, K. P., Perriman, A. W. & Mann, S. Isolation of a highly reactive β -

- sheet-rich intermediate of lysozyme in a solvent-free liquid phase. *J. Phys. Chem. B* **117**, 8400–7 (2013).
78. Prabhu, N. & Sharp, K. Protein-solvent interactions. *Chem. Rev.* **106**, 1616–1623 (2006).
79. Vert, M.; Doi, Y.; Hellwich, K.; Hess, M.; Hodge, P.; Kubisa, P.; Rinaudo, M.; Schue, F. Terminology for biorelated polymers and applications (IUPAC Recommendations 2012). *Pure Appl. Chem.* **84**, 377–410 (2012).
80. Zhang, Y., Patil, A. J., Perriman, A. W. & Mann, S. Enhanced catalytic activity in organic solvents using molecularly dispersed haemoglobin–polymer surfactant constructs. *Chem. Commun.* **49**, 9561 (2013).
81. Kelley, D. & McClements, D. J. Interactions of bovine serum albumin with ionic surfactants in aqueous solutions. *Food Hydrocoll.* **17**, 73–85 (2003).
82. Bio-rad. Protein electrophoresis methods. at <<http://www.bio-rad.com/en-uk/applications-technologies/protein-electrophoresis-methods>>
83. Noda, S., Kamiya, N., Goto, M. & Nakashio, F. Enzymatic polymerization catalyzed by surfactant-coated lipases in organic media. *Biotechnol. Lett.* **19**, 307–309 (1997).
84. Sharma, K. P., Collins, A. M., Perriman, A. W. & Mann, S. Enzymatically active self-standing protein-polymer surfactant films prepared by hierarchical self-assembly. *Adv. Mater.* **25**, 2005–10 (2013).
85. Sharma, K. P. *et al.* Self-organization of glucose oxidase-polymer surfactant nanoconstructs in solvent-free soft solids and liquids. *J. Phys. Chem. B* **118**, 11573–80 (2014).
86. Gallat, F.-X. *et al.* A polymer surfactant corona dynamically replaces water in solvent-free protein liquids and ensures macromolecular flexibility and activity. *J. Am. Chem. Soc.* **134**, 13168–71 (2012).
87. Brogan, A. P. S., Sessions, R. B., Perriman, A. W. & Mann, S. Molecular Dynamics Simulations Reveal a Dielectric-Responsive Coronal Structure in Protein–Polymer Surfactant Hybrid Nanoconstructs. (2014).
88. Brogan, A. P. S. & Hallett, J. P. Solubilizing and Stabilizing Proteins in Anhydrous Ionic Liquids through Formation of Protein-Polymer Surfactant Nanoconstructs. *J. Am. Chem. Soc.* **138**, 4494–4501 (2016).
89. Brogan, A. P. S., Bui-Le, L. & Hallett, J. P. Non-aqueous homogenous biocatalytic conversion of polysaccharides in ionic liquids using chemically modified glucosidase. *Nat. Chem.* **10**, 859–865 (2018).

-
90. Smiglak, M. *et al.* Ionic liquids for energy, materials, and medicine. *Chem. Commun.* **50**, 9228–9250 (2014).
 91. Stroud, R. M. A family of protein cutting proteins.pdf. *Sci. Am.* **231**, 74–88 (1974).
 92. Neurath, H. & Walsh, K. a. Role of proteolytic enzymes in biological regulation (a review). *Proc. Natl. Acad. Sci. U. S. A.* **73**, 3825–32 (1976).
 93. Bordusa, F. Proteases in organic synthesis. *Chem. Rev.* **102**, 4817–68 (2002).
 94. Bongers, J. & Heimer, E. P. Recent applications of enzymatic peptide synthesis. *Peptides* **15**, 183–93 (1994).
 95. Spiro, R. G. Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. *Glycobiology* **12**, 43R–56R (2002).
 96. Patel, T., Gores, G. J. & Kaufmann, S. H. The role of proteases during apoptosis. *FASEB J.* **10**, 587–97 (1996).
 97. Molecular Biotechnological Aspects of Microbial Proteases.pdf. *Microbiol. Mol. Biol. Rev.* **62**, 597–635 (1998).
 98. Raju, R. M., Goldberg, A. L. & Rubin, E. J. Bacterial proteolytic complexes as therapeutic targets. *Nat. Rev. Drug Discov.* **11**, 777–89 (2012).
 99. Gersch, M., Kolb, R., Alte, F., Groll, M. & Sieber, S. A. Disruption of Oligomerization and Dehydroalanine Formation as Mechanisms for ClpP Protease Inhibition. *J. Am. Chem. Soc.* **136**, 1360–1366 (2014).
 100. Bikard, D. *et al.* Exploiting CRISPR-Cas nucleases to produce sequence-specific antimicrobials. *Nat. Biotechnol.* **32**, 1146–1150 (2014).
 101. Lundqvist, M. *et al.* Nanoparticle size and surface properties determine the protein corona with possible implications for biological impacts. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 14265–70 (2008).
 102. Koohmaraie, M. The role of endogenous proteases in meat tenderness.pdf.
 103. Baker, K. C., Taylor, M. A. J., Cummings, N. J., Worboys, K. A. & Connerton, I. F. Autocatalytic processing of pro-papaya proteinase IV is prevented by crowding of the active-site cleft. *Protein Eng.* **9**, 525–529 (1996).
 104. Rawlings, N. D. & Barrett, A. J. Evolutionary families of peptidases. **218**, 205–218
 105. Berg, J. M, Tymoczko, J. L & Stryer, L. in *Biochemsitry* (W H Freeman, 2002).
 106. Kraut, J. Serine Proteases: Structure and Mechanism of Catalysis. *Annu. Rev. Biochem.* **46**, 331–

- 358 (1977).
107. Ekici, O. D., Paetzel, M. & Dalbey, R. E. Unconventional serine proteases: variations on the catalytic Ser/His/Asp triad configuration. *Protein Sci.* **17**, 2023–37 (2008).
 108. Buller, A. R. & Townsend, C. a. Intrinsic evolutionary constraints on protease structure, enzyme acylation, and the identity of the catalytic triad. *Proc. Natl. Acad. Sci. U. S. A.* **110**, E653-61 (2013).
 109. Active, T., Bauer, C., Thompson, R. C. & Blout, E. R. The active centers of *Streptomyces griseus* protease 3, α -chymotrypsin, and elastase: enzyme-substrate interactions close to the scissile bond. *Biochemistry* **15**, 1296–1299
 110. Carrington, J. C. & Dougherty, W. G. A viral cleavage site cassette: identification of amino acid sequences required for tobacco etch virus polyprotein processing. *Proc. Natl. Acad. Sci.* **85**, 3391–3395 (1988).
 111. Bryan, P., Pantoliano, M. W., Quill, S. G., Hsiao, H. & Poulos, T. Site-directed mutagenesis and the role of the oxyanion hole in subtilisin. *Proc. Natl. Acad. Sci. USA* **83**, 3743–3745
 112. Dodson, G. & Wlodawer, A. Catalytic triads and their relatives. *Trends Biochem. Sci.* **23**, 347–52 (1998).
 113. Pelmeshnikov, V., Blomberg, M. R. a & Siegbahn, P. E. M. A theoretical study of the mechanism for peptide hydrolysis by thermolysin. *J. Biol. Inorg. Chem.* **7**, 284–98 (2002).
 114. A. Warshel. How Do Serine Proteases really work. **28**, (1989).
 115. Dissecting the catalytic triad of a serine protease. *Lett. to Nat.* **332**, 564–568 (1988).
 116. Yousef, G. M., Kopolovic, A. D., Elliott, M. B. & Diamandis, E. P. Genomic overview of serine proteases. *Biochem. Biophys. Res. Commun.* **305**, 28–36 (2003).
 117. Jany, K.-D., Lederer, G. & Mayer, B. Amino acid sequence of proteinase K from the mold *Tritirachium album Limber*. *FEBS Lett.* **199**, 139–144 (1986).
 118. Tritirachium, E. Specificity of Proteinase for K from Synthetic Tritirachium Peptides album Limber Kazuyuki MORIHARA and Hiroshige TSUZUKI Shionogi Shionogi & Co., Ltd., Osaka 553 Received Inhibition study . methyl ketone derivatives was determined as follows : **39**, (1975).
 119. Petrotchenko, E. V. *et al.* Use of Proteinase K Nonspecific Digestion for Selective and Comprehensive Identification of Interpeptide Cross-links: Application to Prion Proteins. *Mol. Cell. Proteomics* **11**, M111.013524 (2012).
 120. Betzel, C. *et al.* Structure of a serine protease proteinase K from *Tritirachium album limber* at 0.98 Å resolution. *Biochemistry* **40**, 3080–3088 (2001).

-
121. Betzel, C., Pal, G. P. & Saenger, W. Three-dimensional structure of proteinase K at 0.15-nm resolution. *Eur. J. Biochem.* **178**, 155–171 (1988).
 122. Saenger, W., M. H. W. . Crystal Structure of Calcium-free Proteinase K at 1.5-Å Resolution. **269**, 23108–23111 (1994).
 123. Ebeling, W. *et al.* Proteinase K from *Tritirachium album* Limber. **97**, 91–97 (1974).
 124. Larson, S. B., Day, J. S., Nguyen, C., Cudney, R. & McPherson, A. High-resolution structure of proteinase K cocrystallized with digalacturonic acid. *Acta Crystallographica Section F: Structural Biology and Crystallization Communications* **65**, (International Union of Crystallography, 2009).
 125. Wang, J., Dauter, M. & Dauter, Z. What can be done with a good crystal and an accurate beamline? *Acta Crystallogr., Sect. D* **62**, 1475–1483 (2006).
 126. Masuda, T. *et al.* Atomic resolution structure of serine protease proteinase K at ambient temperature. *Sci. Rep.* **7**, 1–11 (2017).
 127. Patil, A. J. *et al.* Liquid Viruses by nanoscale engineering of capsid surfaces. *Adv. Mater.* **24**, 4557–4563 (2012).
 128. Zhang, Y., Patil, A. J., Perriman, A.W., and Mann, S. Enhanced catalytic activity in organic solvents using molecularly dispersed haemoglobin-polymer surfactant constructs. *Chem. Comm.* **49**, 9561–9563 (2013).
 129. Liu, K. *et al.* Liquefaction of Biopolymers: Solvent-free liquids and liquid crystals from nucleic acids and proteins. *Acc. Chem. Res.* **50**, 1212–1221 (2017).

Chapter 2 Materials and methods

2.1 Introduction

This chapter is divided into two main sections; the techniques and the associated theory for each experimental technique that is central to this thesis, and the details of materials and methods, *i.e.* the source of the chemicals used, the specific machine set-up, and experimental parameters *etc.*

2.2 Techniques and associated theory

2.2.1 Circular dichroism spectroscopy

Circular Dichroism (CD) is a spectroscopic technique that exploits the dichroic nature of optically active materials and can more specifically be utilised to elucidate structure elements of proteins. To this end the technique was used for three distinct purposes; 1. To analyse how each stage of the solvent-free liquid protein synthesis affected the level of secondary structure retention for each protein in question; 2. To ascertain the temperature/structure relationship (a measure of protein stability) of each stage of the solvent-free liquid protein synthesis for each protein in question; 3. To investigate how each stage of the solvent-free liquid protein synthesis alters enzyme (protease) kinetics through the development of CD-based assay.

CD spectroscopy is a technique that measures the differential absorption of electronic transitions within a molecule when left and right circularly polarized light is passed through it¹. Circularly polarized light is a polarization state where the magnitude of the electric field vectors of the light remain constant whilst the direction of the vector rotates about the direction of propagation (k) (Figure 2.1A). When the electric vector is plotted along k , where k is towards the observer, left circularly polarized (LCP) light has an anti-clockwise rotation and right circularly polarised (RCP) light has a clockwise rotation. The combination of two coherent waves over one wavelength where one is LCP and the other is RCP can be represented as a circle about k (Figure 2.1B). If the two circularly polarized waves were to pass through a material and be absorbed differently, the recombined representation would be elliptical, and the emerging light would be elliptically polarized (Figure 2.1C).

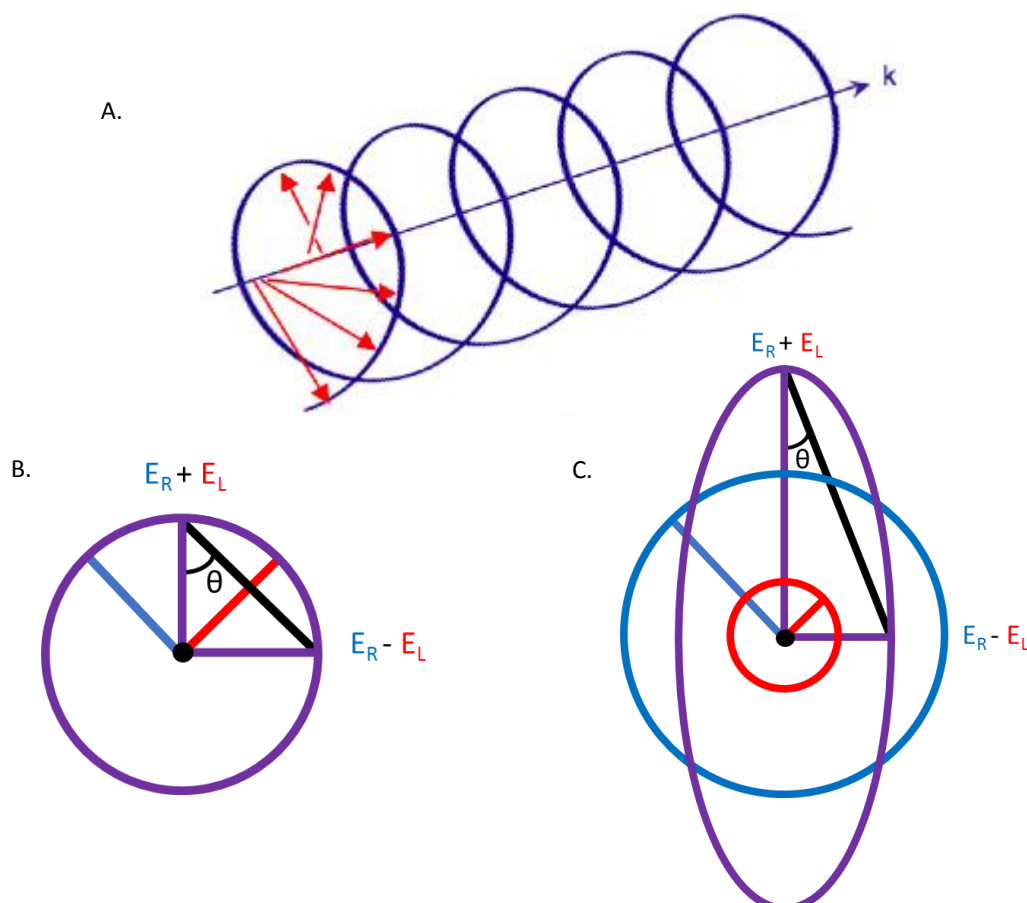


Figure 2.1: **A.** Cartoon depicting right circularly polarized (RCP) light (dark blue line) and the electric field vector of the light (red arrows) with each arrow having the same magnitude and the direction of propagation in direction k , **B.** Cartoon depicting circularly polarized light (purple) composed of equal contributions of RCP-light (blue) and LCP-light (red), and **C.** Cartoon depicting elliptical polarized light (purple) composed of unequal contributions of RCP-light (blue) and LCP-light (red).

It is common place for spectropolarimeters to measure each component individually and not recombine them to form elliptically polarized light, although this is possible. It is this difference in absorption (ΔA) of the left and right components that is measured (Equation 1). Although ΔA is what is measured, historically the measurement is quoted in degrees of ellipticity (θ) (Equation 4), or molar ellipticity ($[\theta]$)^{1,2}. Typically, a CD spectrometer would convert the difference in absorption to molar ellipticity through Equation 5, with a suitable unit for proteins being millidegrees of ellipticity (mdeg).

$$\Delta A = A_L - A_R \quad (1)$$

Equation 1: Where A_L and A_R are the magnitudes of LCP- and RCP-light that is absorbed, respectively.

$$\Delta A = (\varepsilon_L - \varepsilon_R)Cl \quad (2)$$

Equation 2: Where ε_L and ε_R are the molar extinction coefficients for LCP and RCP-light, respectively, C is the molar concentration, and l is the path length (in cm).

$$\Delta\varepsilon = \varepsilon_L - \varepsilon_R \quad (3)$$

Equation 3: Where $\Delta\varepsilon$ is molar circular dichroism value for one specific wavelength.

$$\theta = \tan^{-1} \frac{E_R - E_L}{E_R + E_L} \quad (4)$$

Equation 4: Where θ is the ellipticity of the polarization (in degrees), E_R and E_L are the magnitudes of the electric field vectors of the RCP and LCP-light, respectively.

$$[\theta] = 3298.2\Delta\varepsilon \quad (5)$$

Equation 5: Where $[\theta]$ is the molar ellipticity.

For a molecule to exhibit circular dichroism, having a preference to absorb more LCP-light or RCP-light, the chromophore absorbing the light must either; itself be intrinsically chiral in nature, be associated with a molecule that is chiral in nature, or be able to become chiral if placed within an asymmetric environment. 19 out of the 20 natural amino acids are levoratory and have chiral centres around the α -carbon (C_α), the amide bond absorption within proteins being the chromophore that is probed when using CD spectroscopy to study proteins. The dihedral angles C_α - N bond (ϕ) and C_α - C (ψ) influence the observed electronic transition of this amide bond, the energy and intensity of the transition changing with differing angles, thus being dependent on the secondary structure (Figure 2.2). Common secondary structure motifs found within proteins, *e.g.* α -helices, β -sheets, random coils *etc.*, are known to have specific dihedral angles that are repeated throughout the respective structures, each structure displaying a distinct characteristic far UV CD spectrum.

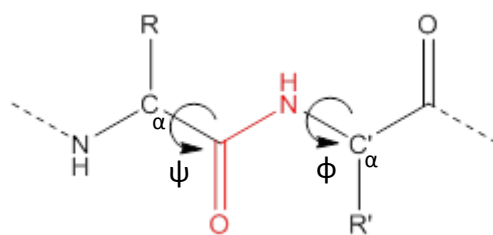


Figure 2.2: Image displaying an amide bond (red) along with the dihedral angles ψ and ϕ associated with the bond.

The electronic transitions that give rise to the CD spectra of proteins are the amide bond transitions $n \rightarrow \pi^*$, which is centred around 220 nm, and the $\pi \rightarrow \pi^*$, which is centred around 190 nm. The $n \rightarrow \pi^*$ transition is often very weak, although it can have large CD bands as it is forbidden electronically but permitted magnetically. In comparison, the $\pi \rightarrow \pi^*$ is a strong transition, not being an electronically forbidden transition. The characteristic CD spectra for an unordered protein (random coil structure) would display a positive peak at 212 nm associated with the parallel $n \rightarrow \pi^*$ transition and a negative peak at 195 nm associated with the $\pi \rightarrow \pi^*$ transition. Both of these transitions are weak when compared to the transitions associated α -helical and β -sheet transitions because the random coil structure has relatively low order. The characteristic CD spectra for a β -sheet structure would display a positive peak at 196 nm associated with the $\pi \rightarrow \pi^*$ transition and a negative peak at 218 nm associated with the $n \rightarrow \pi^*$ transitions. The characteristic CD spectra for an α -helical structure, unlike the previous two structures, displays three characteristic peaks: a red-shifted negative peak at 222 nm that is associated with the $n \rightarrow \pi^*$ transition, and two peaks associated with the $\pi \rightarrow \pi^*$ transition; a positive peak at 192 nm associated with the $\pi \rightarrow \pi^*$ transition that is perpendicular to the axis of the helix, and a negative peak at 209 nm associated with the $\pi \rightarrow \pi^*$ transition that is parallel to the axis of the helix. These two peaks arise due to the strong electric dipole transition moment associated with the $\pi \rightarrow \pi^*$ transition meaning that neighbouring peptide groups interact, a phenomenon called exciton splitting. Interactions along the helical axis give rise to the parallel peak and interactions 90° to the helical axis give rise to the perpendicular peak².

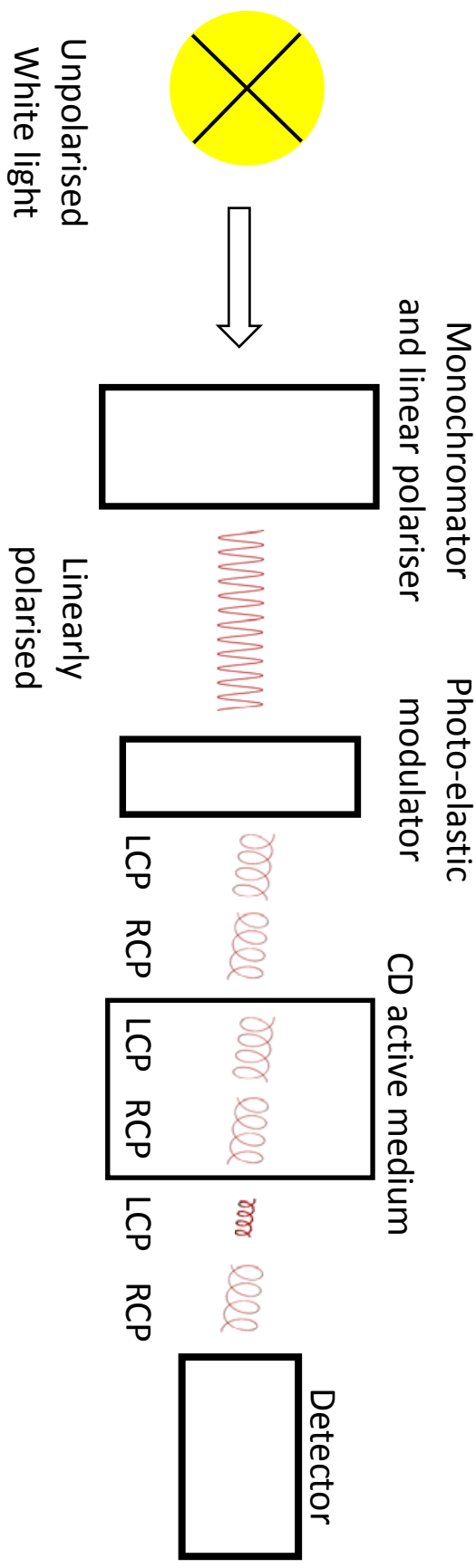
To enable the numerical comparison of multiple CD spectra the concentration of the sample, and therefore the concentration of the chromophore, must be factored, as in Equation 6. For protein-based CD this is the concentration of peptide bonds within the analysed sample, and therefore a peptide-based factor is added, giving rise to mean residue ellipticity ($[\theta]_{mre}$), calculated from ellipticity. The units for $[\theta]_{mre}$ are $\text{deg.cm}^2.\text{dmol}^{-1}$.

$$[\theta]_{mre} = \frac{M\theta}{10(N-1)dc} \quad (6)$$

Equation 6: Where $[\theta]_{mre}$ is the mean residue ellipticity, M is the molar mass of the protein (g.mol^{-1}), θ is the ellipticity (deg), N is the number of residues in the protein (with $N-1$ yielding the number of peptide bonds within the protein), d is the pathlength (cm), and c is the concentration (g.cm^{-3}).

This therefore normalizes the value, yielding one that is independent of chromophore concentration, in the case of proteins the value is independent of protein concentration and peptide length.

Figure 2.3: (Next page) Basic schematic of a circular dichroism spectrometry. The yellow 'X' represents a light source emitting unpolarized white light. This light passes through a monochromator and subsequently a linear polariser which linearly polarises the monochromatic light. The linearly polarised light passes through a photo-elastic modulator that is driven in such a way to circularly polarise the monochromatic light. The circularly polarised light passes through the sample medium, with the left- and right-handed components being absorbed to varying degrees, with the resultant light being detected.



2.2.2 Differential scanning calorimetry

An important technique utilised for the characterization of the solvent-free liquid proteins was differential scanning calorimetry (DSC). This is a thermoanalytical technique used to study the thermal properties of a material, experimentally deriving a relationship between temperature and several material dependent physical properties. A DSC measures the temperature of the system and the heat flow through a sample and reference cell as the temperature increases and decreases at a constant rate. When the sample material passes through a phase transition, the excess heat that is absorbed or radiated due to such a physical transition, relative to the reference sample, is recorded.

The specific instrument used in this work is classified as a conventional hermetic power-compensated DSC, termed due to the instrument's operational mechanism. For this form of DSC, the sample material and a reference material are hermetically sealed within pans, the reference material being one with a well-defined heat capacity (often air). The pans are hermetically sealed in-order to retain a constant mass throughout the heating and cooling ramps of the experiment, even if the sample should evaporate. The two pans are placed onto thermoelectric disks within a furnace and heated/cooled at linear rate. The heat from the furnace is transferred to the samples through the thermoelectric disks, which also monitor the heat flow. Due to the temperatures of phase transitions and heat capacities of the sample and reference material being different, the heat is transferred to the materials at different rates for different temperatures, ensuring there would be no temperature difference between the two cells. In power-compensated DSC the temperature of each cell is kept constant throughout the experiment, the thermoelectric disks either heating or cooling either sample for this to occur. The difference in power being supplied to the sample cell relative to the reference cell is plotted against temperature.

Phase transitions, often referring to transitions between solids, liquids, and gases, occur at material specific conditions when one or more external conditions within the sample's environment is altered, *e.g.* temperature, pressure, *etc.* When a phase transition occurs, certain thermodynamic variables become discontinuous. Taking DSC as an example, phase transitions occur at specific temperatures and manifest as one thermoelectric disk requiring either a higher or lower relative power input to the sample or reference cell to keep them at a constant relative temperature. Depending on whether the phase transition requires a higher or lower relative power input determines whether it is an exothermic or endothermic transition, respectively. Thought of in terms of Gibbs free energy, two or more phases in equilibrium, *e.g.* the triple point of water, ice, and vapour at 273.16 K, have equal Gibbs free energy. A phase transition relates to a global ordering or disordering of a sample, *i.e.* the interparticle interactions within a system alter and allow either a higher or lower degree of freedom. This therefore means that the entropy of the system is changing.

A first-order phase transition of a material involves latent heat of transition, either radiating or absorbing a fixed amount of energy per unit volume of the sample. Taking DSC as an example this

involves the thermoelectric disk either requiring a higher or lower power input to the sample cell relative to the reference cell. Throughout this thesis the DSC data displayed always shows a higher relative power input as positive on the DSC curves, with a lower relative power input being displayed as negative on the DSC curves.

A melting transition (T_m) occurs when a material transitions from a solid state to a liquid state. This decrease in entropy necessitates an increase in enthalpy to maintain an equal Gibbs free energy, manifesting as a requiring a higher relative power input to maintain a constant temperature, being an endothermic transition (Figure 2.4). A crystallisation transition (T_c) involves an increase in entropy as a disordered liquid forms a crystal structure, manifesting as requiring a lower relative power input to maintain a constant temperature, being an exothermic transition (Figure 2.4).

The glass-liquid transition (T_g) occurs when a hard and often brittle solid transitions to a viscous liquid state over an increased temperature. This transition occurs in amorphous materials, or materials containing amorphous regions, and is a reversible transition. The glass-liquid transition is where the amorphous material gradually loses degrees of freedom, leading to a gradual increase in entropy. This gradual increase in entropy is in contrast to first-order transitions that involve a discontinuity and a sharp change in entropy and is therefore not considered a true phase transition. In terms of DSC, this manifests as a gradual increase in relative power input to the sample cell, being an endothermic transition. The glass-liquid transition is always lower than the corresponding melting transitions temperature, if the material possesses one, and the quoted temperature is the midpoint on the curve (Figure 2.4).

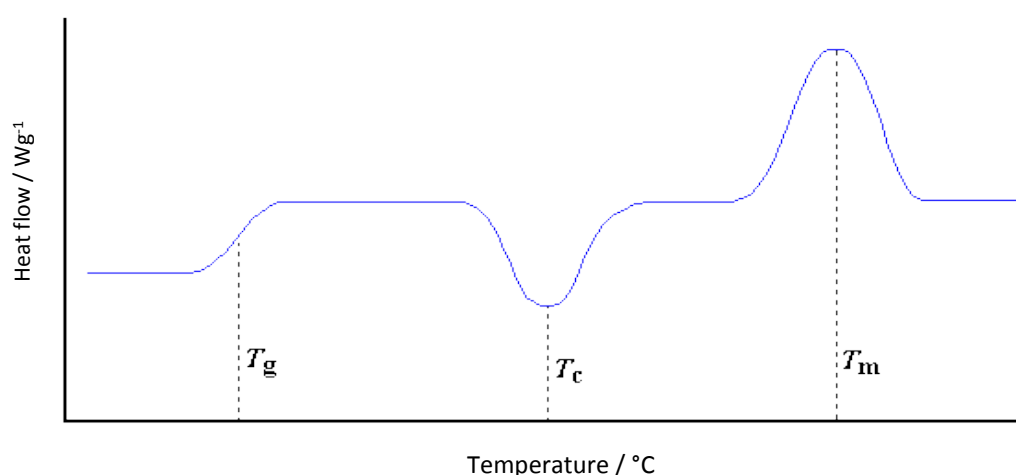
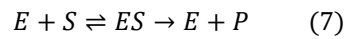


Figure 2.4: A representative DSC curve, heat flow (Wg^{-1}) against temperature ($^{\circ}\text{C}$). An endothermic glass-liquid transition (T_g), endothermic melting transition (T_m), and exothermic crystallization transition (T_c) are displayed. It should be noted that a DSC trace can involve heating and cooling cycles, revealing thermal hysteresis in a material. It should also be noted that the direction of the peak, *i.e.* up or down, is dependent on each individual DSC apparatus.

2.2.3 Michaelis-Menten kinetics

Michaelis–Menten kinetics is a simple well used model for enzyme kinetics that describes the rate of an enzymatic reaction through the relationship between the reaction rate of product formation ($[P]$) to the concentration of the substrate ($[S]$) (Equation 7). More precisely substrate (S) reversibly binds to an enzyme (E) forming an enzyme-substrate complex (ES), which leads to the irreversible reaction between E and S to form a product (P) along with the free enzyme (E) (Equation 7)³.



Equation 7: Enzyme catalysis, where E is enzyme, S is substrate, ES is the enzyme-substrate complex, and P is the product.

The Michaelis–Menten equation describing the reaction rate (v) for a system such as this is:

$$v = \frac{V_{max}[S]}{K_M + [S]} \quad (8)$$

Equation 8: Michaelis–Menten equation where V_{max} is the maximum reaction velocity of the system when substrate is at saturation concentrations, K_M is the Michaelis–Menten constant (the concentration of the substrate when the reaction velocity is 50% V_{max}), and $[S]$ is the concentration of the substrate, example plot in Figure 2.5.

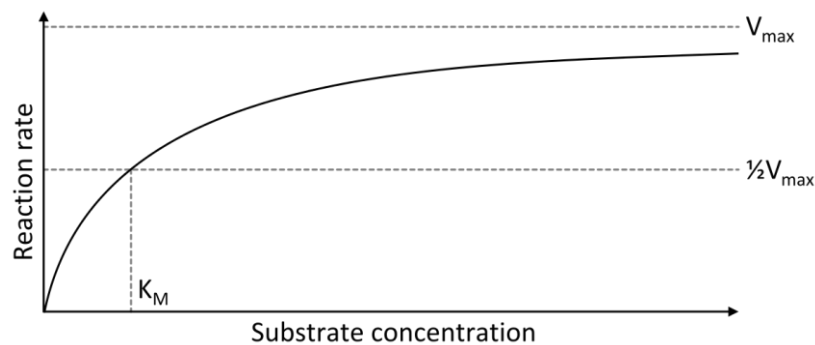
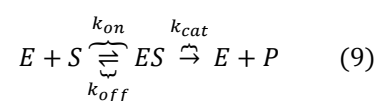


Figure 2.5: Example of a Michaelis–Menten plot, displaying V_{max} and K_M , the maximum enzymatic rate and the Michaelis–Menten constant respectively. $1/2V_{max}$ is the value of the rate at a substrate concentration of K_M .

A simplified derivation of the Michaelis–Menten equation, provided by Briggs and Haldance⁴, is as follows:

Starting with an extended relationship equation:



where k_{on} is the bimolecular association rate constant of enzyme-substrate binding, k_{off} is the unimolecular rate constant of the ES complex dissociating to regenerate free enzyme and substrate; and k_{cat} is the unimolecular rate constant of the ES complex dissociating to give free enzyme and product P .

Now that the rate constants are defined, rate of change equations can be written for each of the chemical species defined, however the rate of change equation for the ES complex is the most important here:

$$\frac{d[ES]}{dt} = k_{on}[E][S] - (k_{off} + k_{cat})[ES] \quad (10)$$

This describes how the concentration of $[ES]$ changes over time in relation to the binding and de-binding rate constants, k_{on} and k_{off} respectively, along with the rate constant for the formation of product P , being k_{cat} . The steady-state approximation can be applied to this equation due to the assumption that the initial ES concentration in an enzyme system rapidly approaches a steady-state, which holds true for the majority of enzyme systems. The steady-state approximation only applies to the initial rate of the reaction. Therefore,

$$\frac{d[ES]}{dt} = 0 \quad (11)$$

$$k_{on}[E][S] = (k_{off} + k_{cat})[ES] \quad (12)$$

The concentration of free enzyme $[E]$ is the total enzyme concentration $[E_T]$ minus $[ES]$. Substituting this and rearranging equation ... we can calculate $[ES]$:

$$k_{cat}[ES] = \frac{k_{cat}[E_T][S]}{\left(\frac{k_{off} + k_{cat}}{k_{on}}\right) + [S]} \quad (13)$$

$$v = k_{cat}[ES] \quad (14)$$

From this, and knowing that $[ES]=[E_T]$ when at maximum reaction velocity (V_{max}), we arrive at the Michaelis–Menten equation:

$$V_{max} = k_{cat}[E_T] \quad (15)$$

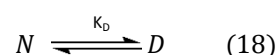
$$K_M = \frac{k_{off} + k_{cat}}{k_{on}} \quad (16)$$

$$v = \frac{V_{max}[S]}{K_M + [S]} \quad (17)$$

2.2.4 Calculation of Gibbs free energy

The following discusses the calculations for Gibbs free energy of denaturation as detailed by Dr. Alex Brogan in his thesis and subsequent papers⁵.

Using circular dichroism (CD), thermal denaturation of a protein was monitored and used within an equilibrium study. A two-state model of equilibrium denaturation was used due to the denaturation pathway of a protein being approximately two-state in nature and permitted the calculation of thermodynamic values without the requirement of thermochemical measuring. This model simply displays protein denaturation as moving from a native (N) state to a denatured (D) state.



where K_D is the equilibrium constant for denaturation. Using the method as described by Pace⁶, the intensity of the 222nm circular dichroic peak obtained was used as an order parameter (y) to express the fraction denatured (f_D),

$$f_D = \frac{(y - y_N)}{(y_D - y_N)} \quad (19)$$

where y_N is the value of the observable order parameter for the native state, and y_D is the value for the denatured state. In order to improve the calculation of f_D the baselines pre- and post- transition were corrected with respect to temperature (T)⁷ using linear expressions for both y_N and y_D ,

$$y_N = a_N + b_N T \quad (20)$$

$$y_D = a_D + b_D T \quad (21)$$

where a_N , b_N , a_D , and b_D are temperature independent variables. The calculated fraction denatured (Equation 19) was then used to calculate the equilibrium constant for unfolding (K_D),

$$K_D = \frac{f_D}{1 - f_D} \quad (22)$$

which was in turn then used to calculate the Gibbs free energy of denaturation (ΔG_D),

$$\Delta G_D = -RT \ln K_D \quad (23)$$

where R is the ideal gas constant and T is temperature. ΔG_D was plotted against temperature (Equation 23), with a linear fit being added to the transition region; the region where ΔG_D transitions from being positive to being negative. This linear fit describes the linear section of the ΔG_D plot around the denaturation temperature (T_m) in the range $+5 \text{ kJ.mol}^{-1} \geq \Delta G_D \geq -5 \text{ kJ.mol}^{-1}$. At T_m , $\Delta G_D = 0$ allowing for the calculation of the entropy of denaturation (ΔS_m) and the enthalpy of denaturation (ΔH_m) from the gradient and intercept (Equation 24), respectively, of the linear fit:

$$\Delta G_D = 0 = \Delta H_m - T_m \Delta S_m \quad (24)$$

2.3 Materials and methods

2.3.1 Materials

Proteinase K from *Tritrachium albumin* (P6556) (PDB:2PKC) was purchased from Sigma-Aldrich as a lyophilised powder (≥ 30 units/mg protein) and *bovine serum albumin* (BSA) (PDB:4F5S) was purchased from Thermo Fisher Scientific, both were used as supplied. Protein concentrations were determined using a Lambda 25 Perkin-Elmer UV-Vis spectrometer and an Implen P300 UV-Vis nanophotometer. 3-(Dimethylamino)-1-propylamine (DMAPA; D145009), N-(3-dimethylaminopropyl)-N'-ethylcarbodiimide hydrochloride (EDC; E1769), and the surfactant poly(ethylene glycol) 4-nonylphenyl-3-sulfopropyl ether ($\text{C}_9\text{H}_{19}\text{-C}_6\text{H}_4\text{-(OCH}_2\text{CH}_2\text{)}_{20}\text{O(CH}_2\text{)}_3\text{SO}_3^-\text{K}^+$; $\text{S}=[\text{C}_9\text{PhE}_{20}\text{OPrSO}_3]^-$; $\text{Mr}=1197$; 473197) (S_1), were used as supplied by Sigma-Aldrich. Brij-L23 (Polyoxyethylene (23) lauryl ether: $\text{CH}_3(\text{CH}_2)_{10}\text{CH}_2(\text{OCH}_2\text{CH}_2)_n\text{OH}$; $n=23$; $\text{Mr}=1274$; 16005) was purchased from Sigma-Aldrich and oxidized using a TEMPO-mediated oxidation protocol, synthesising S_7 .

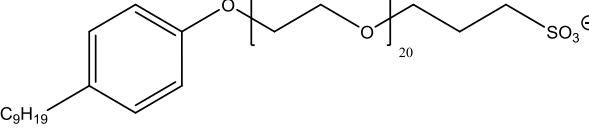
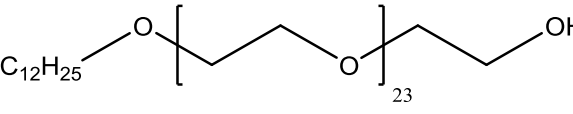
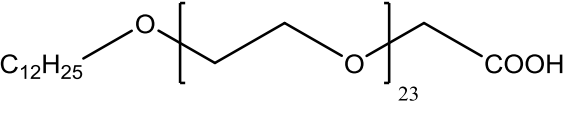
2.3.1.1 Brij-L23-COOH

The following reagents and solvents were purchased and used as received. Polyoxyethylene (23) lauryl ether (Brij-L23) ($1260.2 \text{ g.mol}^{-1}$), 2,2,6,6-tetramethyl-1-piperidinyloxy (TEMPO), sodium bromide (NaBr), sodium hypochlorite (10-15 % available chlorine) and chloroform (CHCl_3) were all purchased from Sigma Aldrich. The experimental procedure was based on previous literature⁸, described briefly below.

Brij-L23 (2g) was dissolved in 50 mL of deionised water and TEMPO (30 mg), sodium bromide (50 mg) and sodium hypochlorite solution (5 mL). The solution was adjusted to pH 11 using concentrated NaOH and then stirred for 24 hours before the reaction was quenched with the addition of ethanol (10 mL). The solution was then adjusted to pH 1 using concentrated HCl, before being extracted with chloroform (3x 80 mL) and dried through rotary evaporation. The resulting white solid was re-dissolved in ethanol (45 mL), and then recrystallized overnight at -20°C . The product was filtered, re-dissolved in ethanol (20 mL) and the solvent removed under reduced pressure to yield a waxy white solid, BrijL23-COOH at 70% yield.

The hydroxy terminus of Brij-L23 (Table 2.1) was oxidised *via* a TEMPO mediated oxidation to yield carboxylated polyoxyethylene (23) lauryl ether (Brij-L23-COOH). The modified surfactant was purified by solvent extraction and multiple re-crystallisations. The synthesis was adapted from previous literature⁸, with the final product being analysed through nuclear magnetic resonance (NMR) spectroscopy.

Table 2.1: Structures and molecular weights of poly(ethylene glycol) 4-nonylphenyl-3-sulfopropyl ether, Brij-23, and modified Brij-L23-COOH.

ID	Name	Structure	M / g.mol ⁻¹	PDI
S1	Poly(ethylene glycol) 4-nonylphenyl-3-sulfopropyl ether		1197.7	1.03
	Polyoxyethylene (23) lauryl ether		1260.2	1.07
S7	Carboxylated polyoxyethylene (23) lauryl ether		1274.2	1.07

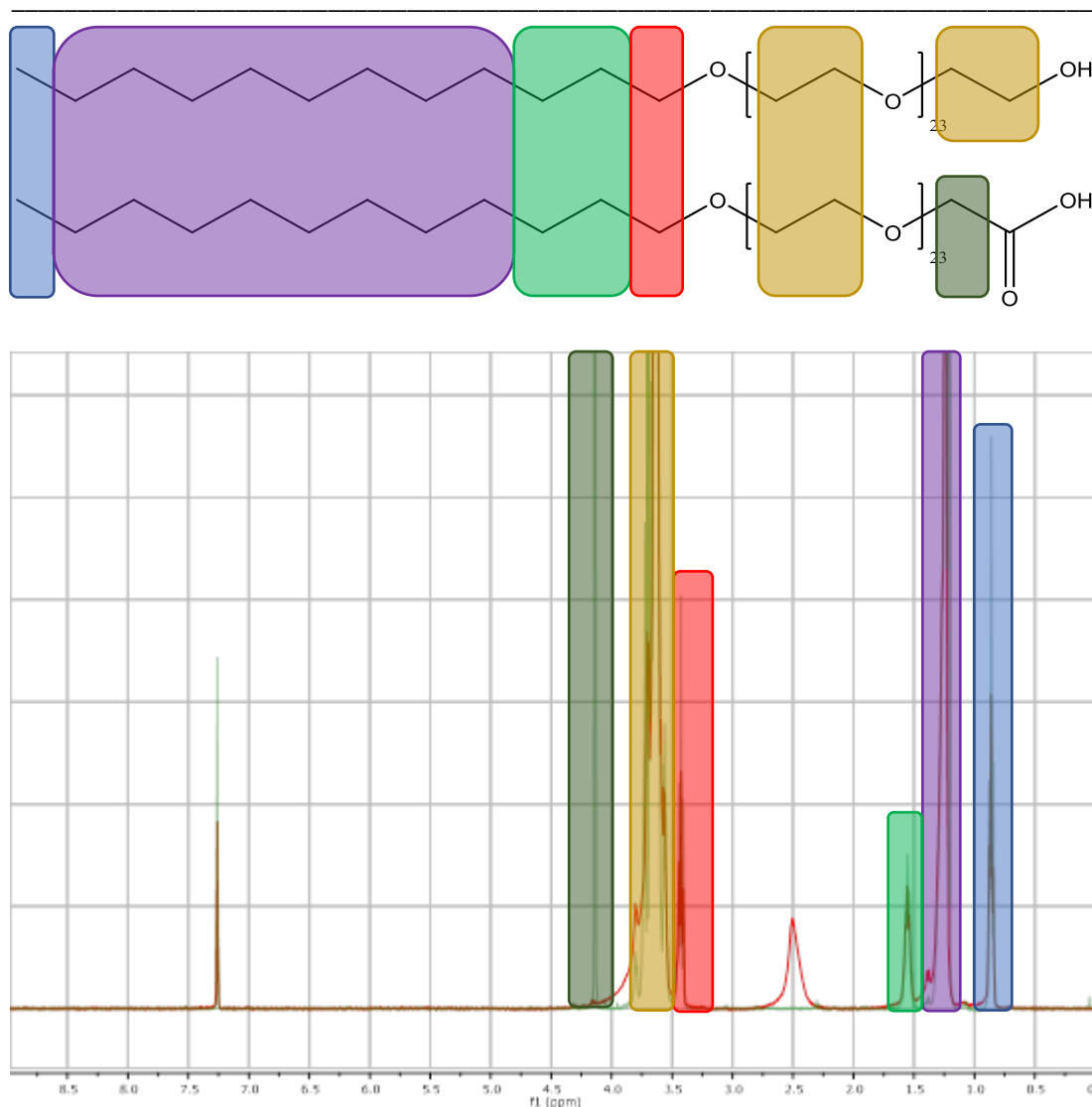


Figure 2.6: **Top structure.** Polyoxyethylene (23) lauryl ether, **Bottom structure.** Carboxylated Polyoxyethylene (23) lauryl ether. NMR spectra of Polyoxyethylene (23) lauryl ether (**red**) and Carboxylated Polyoxyethylene (23) lauryl ether (**green**), with the coloured boxes linking the peaks in the spectra to the proton environments within the relative structures.

^1H NMR spectra (Figure 2.6) were obtained using a Varian 400 MHz spectrometer. ^1H NMR spectra are reported as δ in units of parts per million (ppm) relative to chloroform (δ 7.26, s). Multiplicities are reported as follows: s (singlet), d (doublet), t (triplet), q (quartet), p (quintuplet) and m (multiplet). ^1H NMR for polyoxyethylene (23) lauryl ether (Figure 2.6 (**red**)) (400 MHz, CDCl_3) 3.81-3.54 (m), 3.46-3.41 (q), 1.57-1.54 (q), 1.27-1.24 (q), 0.88-0.84 (m). ^1H NMR for carboxylated polyoxyethylene (23) lauryl ether (Figure 2.6 (**green**)) (400 MHz, CDCl_3) 4.14 (s), 3.80-3.54 (q), 3.46-3.41 (q), 1.57-1.52 (q), 1.27-1.20 (q), 0.88-0.84 (m). The broad singlet at 2.5 in the reagent spectrum (**red**) is from an O-H (alcohol) environment, and it is not seen in product (**green**) spectrum because occasionally the proton can exchange with the deuterium in the deuterated solvent. The spectra are comparable to previous literature⁵.

The appearance of the singlet at 4.14 is indicative of a new proton environment post-reaction. As the peak is a singlet, this means there are no proton environments adjacent to it, this not being found in the chemical structure of the starting reagent. The singlet peak integrates to 2 meaning there are two

protons in this environment, as seen in the chemical structure. It is also at the correct chemical shift for the expected sp^2 hybridised environment, being closer to 5 than the other sp^2 hybridised environments as it is proximal to a double bonded oxygen.

2.3.2 Methods

2.3.2.1 General synthesis of solvent-free liquid proteins

2.3.2.1.1 Protein cationization

External aspartic and glutamic acid residues are coupled with 3-(dimethylamino)-1-propylamine (DMAPA) (Figure 2.7) through carbodiimide activation (Figure 2.8)⁵. Solutions of DMAPA were adjusted to pH 5.5 using 6M HCl and were added dropwise to a stirred protein solution; either proteinase K or BSA. The coupling reaction was initiated by the addition of N-(3-dimethylaminopropyl)-N'-ethylcarbodiimide hydrochloride (EDC) immediately and every hour for the following 3 hrs. The solution was stirred for 12 hrs, with a pH value of 5.5 being maintained throughout using dilute HCL. The solutions were then centrifuged (8000 rpm) to remove any precipitate, with the supernatant then being dialysed (Visking dialysis tubing, 12–14 kDa molecular weight cut off (MWCO)) extensively against Milli-Q quality water for 48 h to produce stable solutions of DMAPA-cationized proteinase K (C-PK) and DMAPA-cationized BSA (C-BSA).

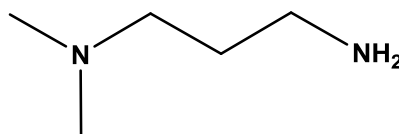


Figure 2.7: Structure of 3-(Dimethylamino)-1-propylamine (DMPA).

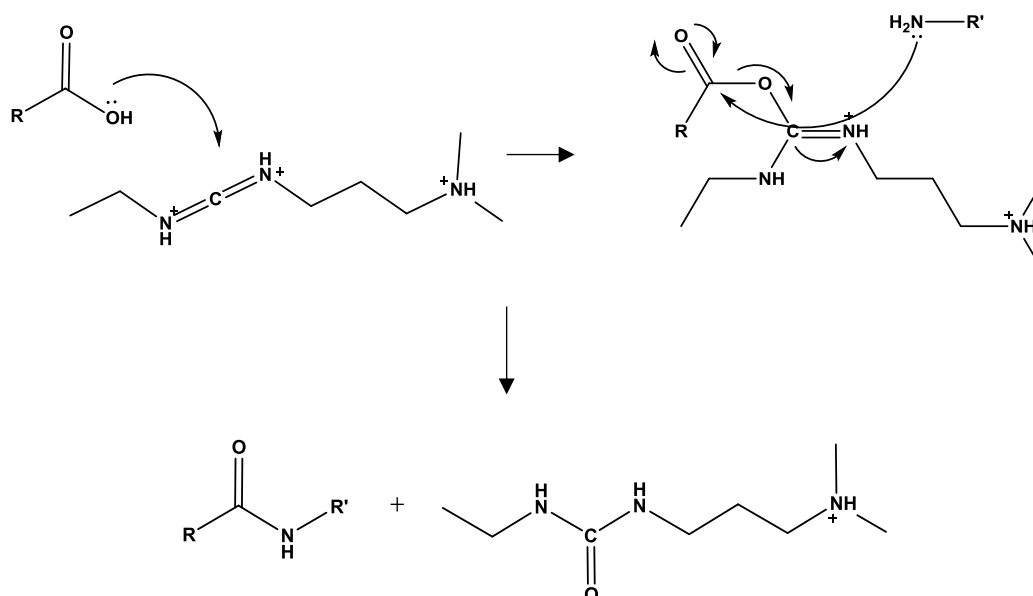


Figure 2.8: Reaction scheme displaying the EDC reaction mechanism, where the carboxylic acid is coupled to an amine compound through activation by N-(3-dimethylaminopropyl)-N'-ethylcarbodiimide hydrochloride (EDC).

2.3.2.1.2 Protein polymer surfactant nanoconjugates

Conjugation of the anionic polymer surfactants S_1 and S_7 to C-PK and C-BSA was carried out by adding an aqueous solution of stated cationized proteins ($4 \text{ mg} \cdot \text{mL}^{-1}$) to a stirred viscous liquid of the diluted surfactant to give a protein:surfactant ratio of 1:1 for C-BSA and C-PK. The solution was subsequently stirred for 12 hr followed by centrifugation (8000 rpm) to remove any precipitate, with the supernatant being dialysed (Visking dialysis tubing, 12–14 kDa MWCO) extensively against Milli-Q quality water for 48 hr to produce stable aqueous solutions of nanoconjugates, designated $[\text{C-BSA}][S_x]$ and $[\text{C-PK}][S_x]$ ($x = 1$ or 7).

2.3.2.1.3 Protein melts

Aqueous samples of protein-surfactant nanoconjugates were lyophilized for 48 hrs and the resulting nanostructured ionic solids were annealed at $\approx 50^\circ\text{C}$ producing clear a viscous pale-yellow liquid of $[\text{C-BSA}][S_x]$ and a clear viscous liquid of $[\text{C-PK}][S_x]$ ($x = 1$ or 7). DSC was utilised to reveal the thermal melting phase transition of a sample. All experiments that involved protein melts were conducted above the thermal transitions identified by DSC experiments to ensure the sample was in the liquid phase as opposed to the solid phase.

2.3.2.1.4 Solvent-free liquid proteins: Liquid density and protein concentration

The concentration of each solvent-free liquid protein created is known to be different because of differing ratios of surfactant molecules present to protein molecules. The density of the solvent-free liquid protein (ρ_{sflp}) and the concentration of protein within the liquid, were calculated from basic density principles. The density was calculated as the sum of the products of the mass fraction and density of the components.

$$\rho_{sflp} = w_p \rho_p + w_s \rho_s$$

where: w is the mass fraction and ρ_p and ρ_s are the densities of the protein and surfactant respectively.

For this calculation, a one to one stoichiometry was assumed; one surfactant molecule per positively charged surface site, a reasonable assumption based on previous literature^{5,9-11}. The density of the surfactant S_1 was detailed by Sigma, with it being assumed that the modification to form S_7 did not change the density. The density of each protein was taken as the value used by crystallographers (1.35 g.cm⁻³). The calculated densities and protein concentrations are displayed in Table 2.2.

Table 2.2: Table displaying the mass fraction, density, and protein concentration of solvent-free liquid BSA and PK

Protein	Mass fraction (w_p)	Density (ρ_{sflp}) (g.cm ⁻³)	Protein concentration (mg.mL ⁻¹)
[C-BSA][S_1]	0.22	1.07	238
[C-BSA][S_7]	0.21	1.06	227
[C-PK][S_1]	0.39	1.12	434
[C-PK][S_7]	0.37	1.12	417
[C-Myo][S_1]	0.24	1.07	259
[C-Myo][S_7]	0.23	1.07	247

Table 2.3: Full list of all solvent-free liquid proteins used throughout this thesis.

ID	Protein	Surfactant [S_n]
[C-BSA][S_n]	Cationized <i>bovine serum albumin</i>	S_1, S_7
[C-FBSA][S_n]	Fluorescently tagged cationized <i>bovine serum albumin</i>	S_1
[C-PK][S_n]	Cationized proteinase K	S_1, S_7
[C-Mb][S_n]	Cationized myoglobin	S_1, S_7

2.3.2.2 Fluorescent tagging

2.3.2.2.1 BSA

Fluorescein isothiocyanate (FITC) is a popular amine-reactive fluorescein-based dye synthesised through an isothiocyanate modification of the 5- or 6-carbon position on the lower of the three fluorescein rings. This synthetic modification produces two near-identical isomers, however, due to chemical differences affecting analysis and separation, the 5-carbon derivative is preferred. FITC reacts with and forms stable products with primary amine groups such as solvent exposed lysine and arginine residues often found on a proteins surface. The primary, often charged, amine nucleophilically attacks the electrophilic central isothiocyanate carbon creating an isothiourea linkage in the form of a thiourea bond (Figure 2.9)¹².

The FITC conjugation to BSA is as follows; 50 mg of BSA is dissolved within 10 mL of a carbonate buffer solution (50 mM) (1.68 g NaHCO₃ + 0.4 g Na₂CO₃ + 200 mL H₂O), 0.5 mL NaCl (150 mM) and an amount of FITC (2 mg.mL⁻¹ within Dimethyl sulfoxide (DMSO)). This solution is wrapped to prevent light entering the sample and then placed within a fridge for 12 hrs. Once removed the solution is dialysed,

with a cut off weight of 7kD, against milliQ water for 48 hours with regular changes of water. The mechanism for FITC conjugation is detailed in Figure 2.9.

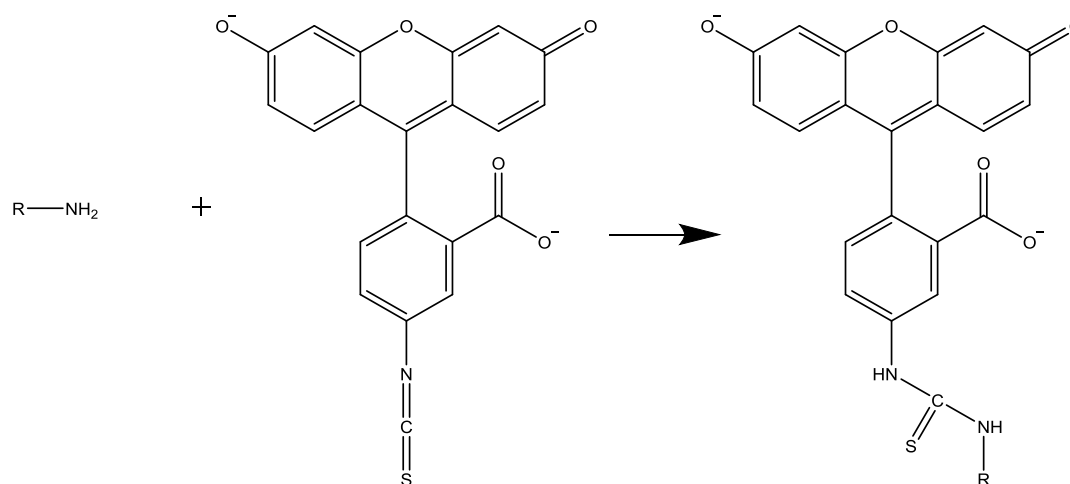


Figure 2.9: A schematic displaying an isothiourea linkage reaction between the amine group (NH_2) of a general amine containing molecule (R) and fluorescein isothiocyanate (FITC) producing an isothiourea linkage between the two molecules.

2.3.2.2.2 Silica

Amphiphilic silica nanoparticles were obtained as a gift from Prof. B. P. Binks, displaying $\approx 50\%$ silanol ($-\text{O}_3\text{SiOH}$) and $\approx 50\%$ dimethylsilane ($-\text{O}_2\text{Si}(\text{CH}_3)_2$) groups (SiO_2_{50}). The two part tagging method of silica nanoparticles with Rhodamine B isothiocyanate (RITC), adapted from work by Li *et al.*¹³, is as follows; amphiphilic silica nanoparticles (20 mg, 20-30 nm) were suspended in ethanol (4 mL) before addition of 3-aminopropyltriethoxysilane (APTES) ($10\ \mu\text{L}\ 10\ \text{mg}\cdot\text{mL}^{-1}$ in ethanol) and triethylamine ($10\ \mu\text{L}\ 5\ \text{mg}\cdot\text{mL}^{-1}$ in ethanol) functionalizing the silica surface with amine groups. It was stirred at room temperature for 24 hours. An ethanoic solution of Rhodamine B isothiocyanate (RITC) was added and stirred at room temperature for a further 5 hours. The suspension was centrifuged at 5000 rpm for 5 minutes and the ethanol layer is removed. Three further ethanol washes were carried out, with the ethanol layer being replaced by dodecane after the final wash. The tagged silica nanoparticles were stored in dodecane until use.

2.3.2.3 Colloidosome formation

The colloidosomes preparation method, adapted from work by Li *et al.*¹³, for the encapsulation of highly concentration protein nano-conjugates is as follows; solvent-free liquid protein (30 mg) was dissolved in water ($50\ \mu\text{L}$) and left for 1 hr to fully dissolve. Dodecane (1.5 mL) was added and the sample was placed within the oven for 30 minutes at $50\ ^\circ\text{C}$. Partially hydrophobised silica nanoparticles (20 mg) were added. When the tagged silica was added to the colloidosome preparation it was added in a 1:10 ratio of tagged to untagged and is termed 'doped silica'. Shearing force was applied by homogenising the sample at 10,000 rpm for 60 s creating a Pickering emulsion. This emulsion was

crosslinked by the addition of Tetramethyl orthosilicate (TMOS) (20 μ L) and placed on a rotary mixer for 5 minutes, subsequently being placed into an oven at 50 °C for 12 hours.

Drying of the colloidosomes was carried out using a method developed by Alex Brogan in unpublished work. The crosslinked colloidosome sample was centrifuged at 2000 rpm for 1 minute before removal of the dodecane. The sample is then placed in an oven at 70 °C for 48 hrs and then placed under vacuum in the presence of phosphorus pentoxide (P_2O_5) for a further 48 hrs.

2.3.2.4 Circular dichroism spectroscopy (CD)

CD experiments were conducted on a Jasco J-810 or J-815 circular dichroism spectropolarimeter between 180-290 nm at a scan speed of 20 nm.min⁻¹, an integration time of 16 seconds, a data interval of 0.1 nm, and with a minimum of 4 repeats. The measurements were temperature controlled using a Jasco PTC-423 peltier control unit that had a temperature range from -40–130 °C. When the temperature of the sample was being changed over time the temperature increase was 0.5 °C.min⁻¹, with a scan every 0.5 °C and each complete set of scans being repeated 3 times. Solution samples were run in 1 mm thick quartz cuvettes, with a typical protein concentration of 0.2 mg.mL⁻¹. Solvent-free liquid samples were run as a thin film spread evenly between two quartz plates at a temperature above the known melting temperature of the solvent-free liquid in question.

2.3.2.4.1 Synchrotron radiation circular dichroism (SRCD)

SRCD experiments were conducted at the B23 beamline at Diamond Light Source in Didcot. Each experiment was run between 180-290 nm at a scan speed of 50 nm.min⁻¹, had an integration time of 2 seconds, and a data interval of 1 nm. The temperature of the experiments was controlled using a Quantum TC 125 temperature controller. Solvent-free liquid samples were run as a thin film spread evenly between two quartz plates.

CD experiments were performed on ten pairings of proteinase K and BSA; N-PK v N-BSA, N-PK v C-BSA, N-PK v aqueous [C-BSA][S_x], [C-PK] v N-BSA, [C-PK] v [C-BSA], [C-PK] v aqueous [C-BSA][S_x], aqueous [C-PK][S_x] v N-BSA, aqueous [C-PK][S_x] v [C-BSA], aqueous [C-PK][S_x] v aqueous [C-BSA][S_x], and [melt PK] v [melt BSA] (N = native) (x = 1 or 7), using circular dichroism to quantify the change in secondary structure due to protease/substrate binding. The pairings were scanned between 190-260 nm at a scan rate of 20 nm.min⁻¹ with an integration time of 8 s. Kinetic enzyme assays were also performed on all nine aqueous pairings, scanning scanned between 190-260 nm at a scan rate of 100 nm.min⁻¹ with an integration time of 8 s over various period of time. The [melt PK] v [melt BSA] activity assays were performed at the Diamond Light Sauce on beamline B23 over a wavelength range of 185 to 260 nm with an integration time of 1 s and with a 1 nm data interval over various time periods, with the temperature being controlled by a modified Linkam thermal stage, the sample cell comprising of two synthetic quartz plates.

2.3.2.5 Data analysis

All data was analysed either through Microsoft Excel 2016, Origin 8, or Igor pro 6.22, unless explicitly stated. Igor Pro 6.22 was used to smooth CD data for analysis, with the Savitzky-Golay algorithm being utilised for this task. All chemical structures displayed were created in ChemDraw Professional 16.0. Protein Data Bank files were downloaded from '<https://www.rcsb.org/pdb/home/home.do>', and imported into and analysed using a space filling model in Visual Molecular Dynamics (VMD) 1.9 from which the number of accessible cationizable sites was determined. Confocal images were analysed using ImageJ and Velocity 6.3 – Cellular imaging and analysis – PerkinElmer.

2.3.2.6 Differential scanning calorimetry (DSC)

DSC experiments were conducted on a TA Instruments Q100 differential scanning calorimeter. The samples were hermetically sealed within aluminium pans, using air as a reference sample unless otherwise stated. A heating/cooling rate of 10 °C.min⁻¹ was used unless otherwise stated.

2.3.2.7 Dynamic light scattering (DLS)

DLS experiments were conducted on a Malvern Instrument Zetasizer with the use of disposable low volume cuvettes. Protein concentrations were 0.5 mg.mL⁻¹ and the measurements were made at a fixed angle of 173°.

2.3.2.8 Mass spectrometry (MS)

MS experiments were conducted on an Applied Biosystems 4700 MALDI TOF (matrix assisted laser desorption/ionization-time of flight-mass) spectrometer in positive ion linear mode. Samples were mixed in a 1:1:1 ratio of A:B:C solutions where: A – three equivalents dihydroxyacetophenone (DHAP:20.3 mg.mL⁻¹ in EtOH) and one equivalent diammonium hydrogen citrate (DAHC: 18 mg.mL⁻¹ in water), B – 2 vol% trifluoroacetic acid solution, and C – 1 mg.mL⁻¹ protein solution¹⁴. The solution was spotted and dried on a ground steel MALDI-TOF-MS plate.

2.3.2.9 Nuclear magnetic resonance spectroscopy (NMR)

¹H NMR spectra were obtained using Varian 400 MHz spectrometer¹H NMR spectra are reported as δ in units of parts per million (ppm) relative to chloroform (δ 7.26, s). Multiplicities are reported as follows: s (singlet), d (doublet), t (triplet), q (quartet), p (quintuplet) and m (multiplet). Coupling

constants are reported as J values in units of Hertz (Hz). The number of protons (n) is reported as nH and based on spectral integration values.

2.3.2.10 Optical microscopy

Optical microscopy experiments were conducted on a Leica DMI 3000B Fluorescence Optical microscope. Images were collected and processed using Leica Applications Suite.

2.3.2.10.1 Polarized light microscopy

For polarized light microscopy the same set as for optical microscopy was used but with the addition of cross-polarizing lens.

2.3.2.10.2 Confocal microscopy

Confocal microscopy was conducted using a Leica SP5-AOBS confocal laser scanning microscope, with a 40x HCX PL APO lens, equipped with 100 mW Ar laser, attached to a Leica DM I6000 inverted epifluorescence microscope.

2.3.2.11 pH measurements

pH measurements were conducted on a Mettler Toledo MP220 pH meter. Prior to each experiment the pH meter was calibrated using standard buffer solutions of pH 4 and 7.

2.3.2.12 Protein secondary structure estimation

The secondary structure of all proteins was estimated through complex deconvolution of CD spectra using the BeStSel service reference. The service works through the inputting of an individual or group of CD spectra with is then run through a highly refined and complex algorithm that outputs and estimation of the proteins secondary structure in the form of percentages for each of; α -helices, β -sheets, turns and random coils. The service also provides a normalized root mean squared deviation for the spectra, a calculation of error for the fit; a higher NRMSD indicating a less reliable fit. For the secondary structure estimates presented a NRMSD no higher than 0.02 was used, a value of ≤ 0.1 being the traditional cut off.

2.3.2.13 SEM

The SEM instrument used was a JSM-IT300 from JEOL; working distance 16 mm, Acc voltage 15.0 kV, BE detector, and a silver (Ag) coating.

2.3.2.14 Thermogravimetric analysis (TGA)

TGA experiments were conducted on a TA Instruments Q500 Thermogravimetric Analyzer. All samples were heated to and held at 80 °C for one hour to remove atmospheric water, followed by heating to and holding at 110 °C as a method to assess the samples water content as a percentage of original sample weight. The sample was then heated to 600 °C to assess the degradation of the sample as a function of temperature. A heating rate of 10 °C.min⁻¹ was used for each heating stage of the experiment. Experiments were conducted on a platinum pan.

2.3.2.15 UV/Vis spectroscopy

2.3.2.15.1 Solution UV/Vis

Solution UV/Vis experiments were conducted on a PerkinElmer Lambda 700 spectrophotometer fitted with a PTP-6 peltier controller. Samples were run in reduced volume black walled 10 mm quartz cuvettes at a scan speed of 240 nm.min⁻¹ between 700-200 nm at 25 °C. Each experiment was repeated a minimum of 3 times and averaged.

2.3.2.15.2 Diffuse reflectance UV/Vis spectrometry (DR-UV/Vis)

DR-UV/Vis experiments were conducted on a PerkinElmer Lambda 700 spectrophotometer fitted with a Labsphere diffuse reflectance integrating sphere accessory. Solvent-free liquid samples were scanned as a thin film spread evenly between two quartz plates, with a scan speed of 240 nm.min⁻¹ between 700-200 nm.

2.3.2.16 Zeta potential

Zeta potential experiments were conducted on a Malvern Zetasizer Nano ZS. Disposable folded capillary cells were used to conduct experiments on samples with a typical concentration of 0.2 – 1 mg.mL⁻¹. Each sample has a minimum of 3 repeats.

2.4 References

1. Greenfield, N. J. Using circular dichroism spectra to estimate protein secondary structure. **1**, 2876–2890 (2009).
2. Kelly, S. M., Jess, T. J. & Price, N. C. How to study proteins by circular dichroism. *Biochim. Biophys. Acta - Proteins Proteomics* **1751**, 119–139 (2005).
3. Johnson, K. A. & Goody, R. S. The original Michaelis constant: Translation of the 1913 Michaelis-Menten Paper. *Biochemistry* **50**, 8264–8269 (2011).
4. Briggs, G. E. & Haldane, J. B. S. A Further Note on the Kinetics of Enzyme Action. *Biochem. J.* **19**, 1037–1038 (1925).
5. Brogan, A. P. S., Siligardi, G., Hussain, R., Perriman, A. W. & Mann, S. Hyper-thermal stability and unprecedented re-folding of solvent-free liquid myoglobin. *Chem. Sci.* **3**, 1839 (2012).
6. Pace CN. The stability of globular proteins. *CRC Crit. Rev. Biochem.* **3**, 1–43 (1975).
7. Ahmad, F. On the estimation of Stability Parameters from Heat-Induced Conformational Transition Curves of Proteins. *J. Iran. Chem. Soc.* **1**, 99–105 (2004).
8. Gallat, F.-X. *et al.* A polymer surfactant corona dynamically replaces water in solvent-free protein liquids and ensures macromolecular flexibility and activity. *J. Am. Chem. Soc.* **134**, 13168–71 (2012).
9. Brogan, A. P. S., Sharma, K. P., Perriman, A. W. & Mann, S. Enzyme activity in liquid lipase melts as a step towards solvent-free biology at 150 °C. *Nat. Commun.* **5**, 6058 (2014).
10. Brogan, A. P. S., Sharma, K. P., Perriman, A. W. & Mann, S. Isolation of a highly reactive β -sheet-rich intermediate of lysozyme in a solvent-free liquid phase. *J. Phys. Chem. B* **117**, 8400–7 (2013).
11. Perriman, A. W., Cölfen, H., Hughes, R. W., Barrie, C. L. & Mann, S. Solvent-free protein liquids and liquid crystals. *Angew. Chemie - Int. Ed.* **48**, 6242–6246 (2009).
12. Hermanson, G. T. in *Bioconjugation techniques* 259–273 (Academic Press, 2013).
13. Li, M., Huang, X., and Mann, S. Spontaneous growth and division in self-reproducing inorganic colloidosomes. *Small*. **10**, 3291–3298 (2014).
14. Hou, J. *et al.* Enhanced MALDI-TOF MS analysis of phosphopeptides using an optimized DHAP/DAHC matrix. *J. Biomed. Biotechnol.* **2010**, 759690 (2010).

Chapter 3 Characterisation of solvent-free liquid proteins and their aqueous precursors

3.1 Introduction

The global aim of this thesis is to demonstrate, for the first time, that protease activity can be demonstrated in the solvent-free liquid state. Accordingly, the primary object in this chapter is to characterise solvent-free protein liquids of the protease (proteinase K) and the substrate, *bovine serum albumen* (BSA). The solvent-free liquid protein synthesis, detailed in Chapter 2.3.2, involves three distinct stages; an initial protein cationization stage forming a highly positively charged protein, a surfactant conjugation stage forming a protein-polymer nanoconjugate, and a final lyophilization and thermal annealing stage that forms the final solvent-free liquid protein (discussed in Chapter 1.4.1. According, it is essential to investigate how each stage of this synthesis affects proteinase K (PK) and *bovine serum albumin* (BSA). As such, the chapter is split into three sections. The first focuses on the cationization of the proteins, with the second focusing on the creation of protein-polymer nanoconjugates. Protein secondary structure assessments and thermal stability studies form the majority of the investigations, through the use of circular dichroism spectroscopy (CD), with particle size and zeta potential also being investigated. The third section focuses on the thermal properties of the solvent-free material, through the techniques of differential scanning calorimetry (DSC) and thermogravimetric analysis (TGA), this providing further information on the material properties of the novel protein liquids. Finally, the secondary structure and thermal stability of the proteins is studied, showing how the formation of the solvent-free liquid phase affects the conformation of the protein.

3.2 Results

3.2.1 Characterisation of cationized proteins

The initial stage of the solvent-free liquid protein synthesis involves the cationization of a protein through the addition of amine groups, in this case addition of DMAPA, to the surface accessible negatively charged carboxylic acid residues. This is carried out through the use of an EDC mediated coupling reaction (see Section 2.3.2 for details). This reaction can often crosslink molecules/proteins together and form aggregates in the system. Aggregation of cationized proteins in this synthesis is undesirable as the aim of cationization is to maximise the number of positive surface point charges, with aggregation reducing the maximal attainable number. To avoid aggregation throughout the cationization process, high concentrations of DMAPA and EDC may be used alongside low concentrations of protein, this maximising the number of covalently surface-attached amines whilst minimising the aggregation of proteins.

In the following section, native *bovine serum albumin* (BSA), native proteinase K (PK), cationized BSA (C-BSA), and cationized PK (C-PK) are characterised to show that the proteins in question were successfully cationized through mass spectrometry (MS) and zeta potential measurements.

3.2.1.1 Mass spectrometry (MS) optimization

Both the reactant concentration and the pH value that the coupling reaction is conducted under has been shown to be a factor that affects the efficiency of protein cationization. Previous optimal protein cationization pHs have been found to range between pH values of 5-6.5¹, showing that the optimal pH value is protein dependent, with the pKa of the carboxylic side chains varying between proteins. Accordingly, the cationization efficiency for a range of pH values was investigated using MS, comparing the molecular ion peak of the native protein to the cationized protein (Equation 1). Where there were broadly distributed peaks displayed on the spectrum, the most intense peak was selected as the molecular ion peak for that spectrum. Using the known number of solvent accessible negatively charged residues present (Z) (Equation 1) yields the cationization efficiency. Each protein was cationized at four pH values of 5, 5.5, 6, and 6.5.

$$\text{Cationization efficiency} = \left(\frac{\frac{MI_{\text{cationized}} - MI_{\text{native}}}{M_{\text{DMAPA}} - M_{\text{water}}}}{Z} \right) \times 100 \quad (1)$$

Equation 1: Where $MI_{\text{cationized}}$ and MI_{native} are the molecular ion peaks of the cationized and native proteins, respectively, and M_{DMAPA} and M_{water} are the molecular masses of DMAPA (102 g.mol⁻¹) and H₂O (18 g.mol⁻¹), respectively, with cationization efficiency being a percentage.

For native BSA the molecular mass was seen to be 66,462 g.mol⁻¹, as displayed by a sharp peak in the MS spectrum (Figure 3.1), and was in agreement with literature values². Within this spectrum, there was also a secondary shoulder at around 67200 g.mol⁻¹ that had a lower intensity (≈ 0.4); this is thought to be due to an impurity in the BSA sample. The number of acidic solvent-accessible surface residues present on BSA was calculated to be 99 (59 glutamic acid residues and 40 aspartate residues)³. For BSA, the number of DMAPA molecules attached to the surface upon cationization was found to be: pH 5 – 95 DMAPA, pH 5.5 – 99 DMAPA, pH 6 – 97 DMAPA, and pH 6.5 – 79 DMAPA. The values and cationization efficiencies shown in Table 3.1 show that at each pH tested there was high cationization efficiency and that cationizing at a pH value of 5.5 was the most efficient, cationizing 100% of the available residues. There is a clear broadening in the peak MS signal once cationized due to the level of cationization being a distribution across the sample.

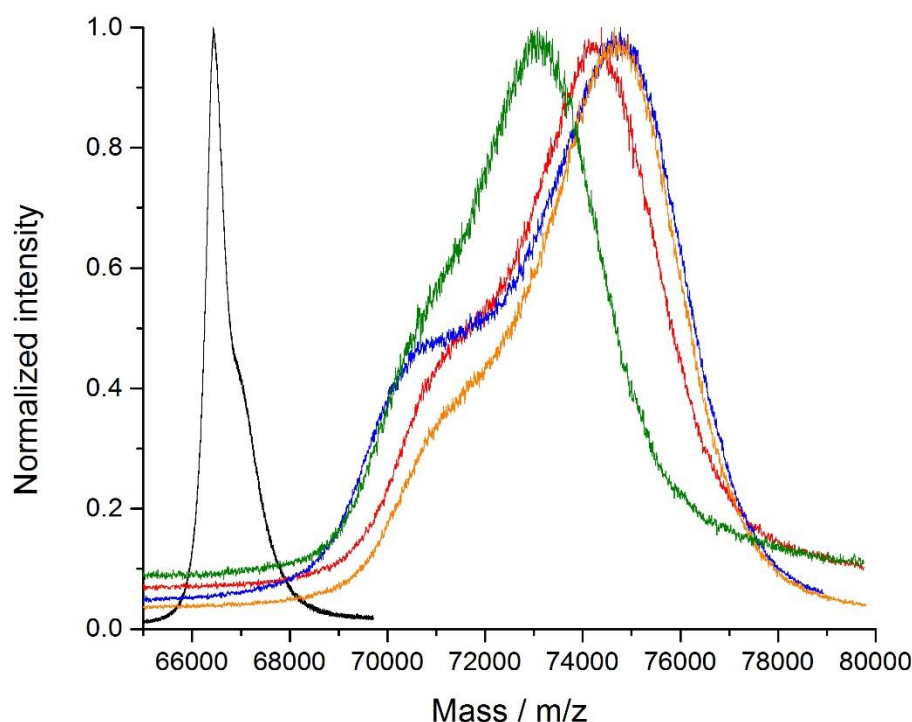


Figure 3.1: MALDI-TOF MS spectra for native-BSA (black), and BSA cationized at pH values of 5 (red), 5.5 (blue), 6 (orange), and 6.5 (green).

Table 3.1: For cationized BSA: cationization pH value, molecular ion peak, number of attached DMAPA molecules, and cationization efficiency.

Sample	Mol. weight (g.mol^{-1})	No. DMAPA molecules	Cationization efficiency (%)
Native-BSA	66,462	NA	NA
C-BSA pH 5	74,387	94	95
C-BSA pH 5.5	74,772	99	100
C-BSA pH 6	74,610	97	98
C-BSA pH 6.5	73,077	79	80

For native-PK the molecular mass was calculated to be $28,993 \text{ g.mol}^{-1}$, as displayed by a sharp peak in the MS spectrum, agreeing with literature values⁴. Within this spectrum, there was also a secondary shoulder at around $29,400 \text{ g.mol}^{-1}$ that had a lower intensity (≈ 0.1); this is thought to be due to an impurity in the PK sample. The number of acidic solvent-accessible surface residues present on PK was calculated to be 18 (5 glutamic acid residues and 13 aspartate residues)⁵. For PK, the number of DMAPA molecules attached to the surface upon cationization was found to be: pH 5 – 14 DMAPA, pH 5.5 – 16 DMAPA, pH 6 – 15 DMAPA, and pH 6.5 – 12 DMAPA. The values and cationization efficiencies shown in Table 3.1 show that at each pH tested there was high cationization efficiency and that cationizing at a pH value of 5.5 was the most efficient, cationizing 89% of the available residues. Again, there is a clear broadening of the peak MS signal once cationized due to the level of cationization being a distribution across the sample.

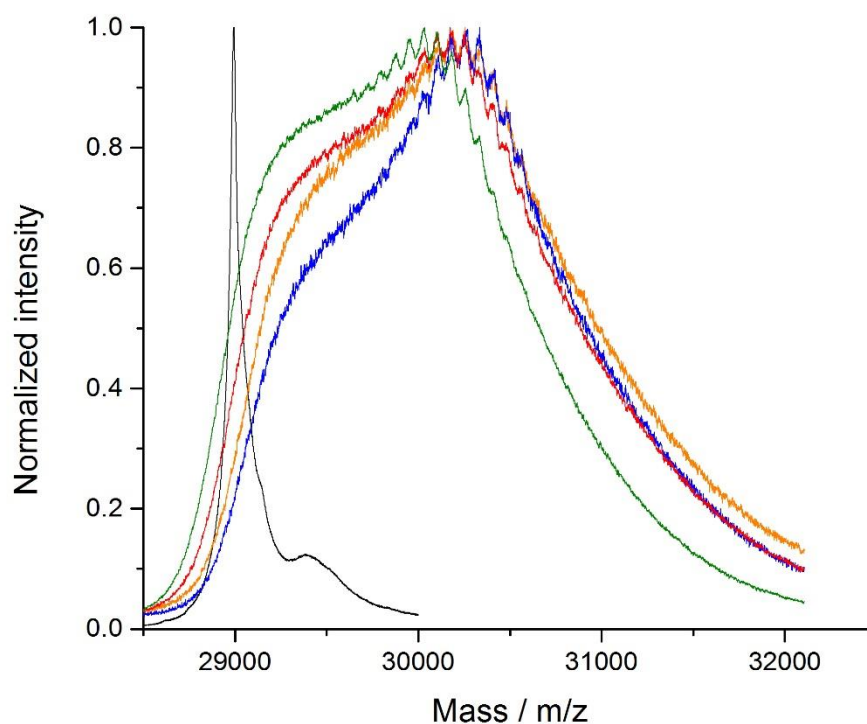


Figure 3.2: MALDI-TOF MS spectra for native-PK (black), and PK cationized at pH values of 5 (red), 5.5 (blue), 6 (orange), and 6.5 (green).

Table 3.2: For cationized PK: cationization pH value, molecular ion peak, number of attached DMAPA molecules, and cationization efficiency.

Sample	Mol. weight (g.mol^{-1})	No. DMAPA molecules	Cationization efficiency (%)
Native-PK	28,993	NA	NA
C-PK pH 5	30,174	14	78
C-PK pH 5.5	30,355	16	89
C-PK pH 6	30,256	15	84
C-PK pH 6.5	30,033	12	69

It is clear from the MS spectra of the various C-BSA and C-PK samples that there is a range of cationization efficiencies present in each sample, displayed as various peaks that are regularly separated by a spacing of $\approx 84 \text{ g.mol}^{-1}$ (Figure 3.3). The spacing between these peaks matches the molecular weight of DMAPA (84 g.mol^{-1}), being the molecule that is being covalently bound to the proteins *via* the EDC reaction. This reaction produces a distribution of proteins with varying numbers of DMAPA molecules attached. As mentioned previously, the most intense and often central peak was chosen to represent a sample due to this being the most prevalent in the sample.

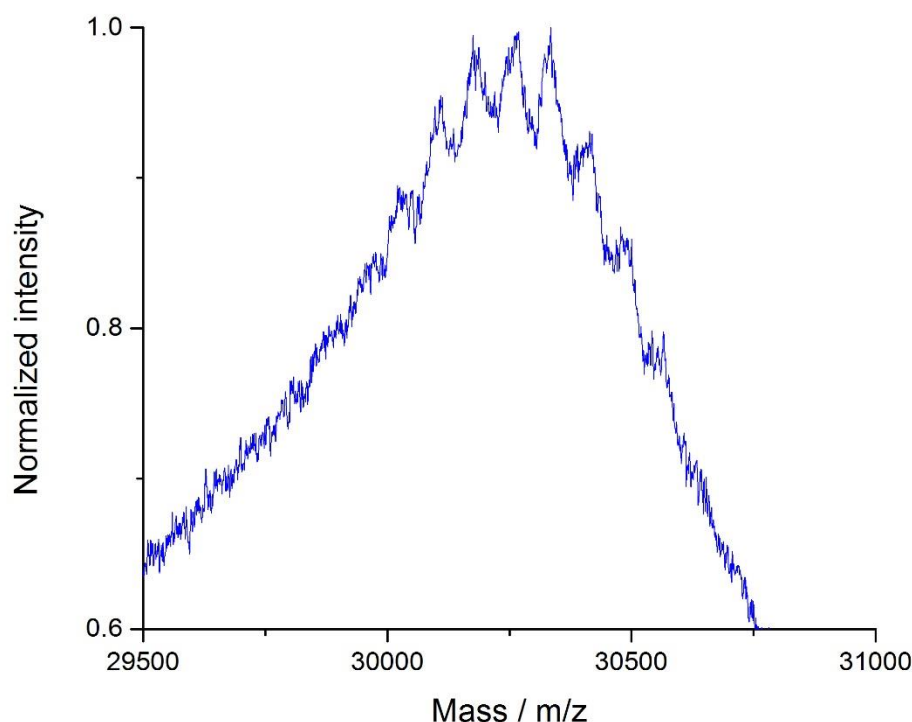


Figure 3.3: Magnified MALDI-TOF MS spectra for cationized PK pH 5.5 (blue) displaying minor peaks with gaps of 84 g.mol^{-1} .

These MS spectra clearly show that both proteins were successfully cationized through the EDC mediated addition of DMAPA molecules, and that both have a maximal cationization efficiency when the EDC reaction was conducted at a pH value of 5.5. For this reason, the cationization pH value of 5.5 was selected for the preparation of all remaining BSA and PK samples reported in this thesis.

3.2.1.2 Charge density measurements using zeta potential

The successful surface cationization of negatively charged sites on a protein has the effect of altering the net surface charge, changing the presented charge from negative to positive and increasing the net surface charge by 2 for each site that is altered. Measuring the zeta potential (ZP) of a colloidal particle, or a particle that can be modelled as such, *e.g.* proteins, can reveal the charge density on the surface, revealing changes from one sample to the next.

The ZP measurements of BSA and C-BSA, conducted in unbuffered water at pH 6 - 6.5, yielded ZP values of $-12 \pm 1 \text{ mV}$ and $10 \pm 3 \text{ mV}$ respectively. The ZP measurements of PK and C-PK, conducted in unbuffered water at pH 6 - 6.5, yielded ZP values of $-15 \pm 2 \text{ mV}$ and $12 \pm 3 \text{ mV}$ respectively. These results provide further evidence for the successful cationization of BSA and PK, with a charge density increase for both proteins once cationized. The sharp peak displayed for native-BSA is indicative of a predominantly monodisperse system, however, the increase in peak number and a large variety of charges displayed by C-BSA indicate the potential for a variety of aggregates forming in the system.

This aggregation could be caused by the partial unfolding of the protein revealing non-polar hydrophobic residues, however, size and structure determination investigations are required to provide further evidence towards this hypothesis.

3.2.1.3 Particle size distribution using dynamic light scattering (DLS)

DLS measurements provide a method to obtain the number particle size distribution throughout a sample. This technique was used to demonstrate the dispersity of the protein systems pre- and post-cationization, along with any subsequent particle size increase or increase in aggregation, either due to crosslinking or to favourable folding. The Zetasizer software creates a mathematical fit, relating the size of the particle to translational diffusion constant and the viscosity of the solvent, that calculates the average radii of hydration within a sample. This means that if any larger aggregates have formed in the system the calculated value could be skewed to be higher than predicted.

3.2.1.3.1 DLS results for BSA and C-BSA

The calculated results for the hydrodynamic radii of BSA and C-BSA were 3.6 nm and 24.4 nm respectively (Figure 3.4). This showed that the cationization of BSA increased the hydrodynamic radius by over 20 nm relative to native-BSA. An increase of this magnitude cannot be attributed merely to the length of the additional amines on the surface of BSA, known to be less than 0.1 nm in length. This effect therefore is attributed to the formation of a new structure, a largely unfolded protein or C-BSA aggregate, which have an average hydrodynamic radius of 24 nm. Aggregates could form through two mechanisms; either the cationization process partially unfolds the protein leading to the exposure of hydrophobic residues and subsequent aggregation, or the presence of highly charged proteins causes aggregation. From previous literature ^{1, 6, 7} it is clear that both cases can occur, therefore a structural study was conducted to investigate the cause of aggregation.

The justification for the use of number particle size distribution comes from a comparison of the number particle size distribution data and the intensity correlation data with previous literature values obtained using small angle neutron scattering (SANS)⁸. The fitting of the autocorrelation function to BSA intensity data yielded hydrodynamic radii of 56 nm, far higher than 3.6 nm obtained from SANS. This value was skewed by the presence of larger aggregates over 100 nm in radii that would dominate the observed scattering.

The fitting of the autocorrelation function to C-BSA intensity data provided supporting evidence for the formation of large C-BSA aggregates, there being a broad range of aggregate sizes of C-BSA up to radii of over 200 nm. The presence of protein aggregates of this size could potentially increase the challenge in forming a solvent-free liquid BSA due to the relatively large size of proteins being the original cause for the lack of liquid state. However, it was clear from the number particle size distribution that the average radius of the aggregated C-BSA was 24 nm, with a low percentage of the

aggregates being of the hundreds of nanometres size (<4%). This average radius of 24 nm is only twice that of the protein ferritin, the first protein to be formed into a solvent-free liquid protein⁹. This aggregation, therefore, might not cause complications in the formation of a solvent-free liquid BSA.

It has also been previously reported that the addition of surfactant molecules to the surface of a cationized protein can lead to the reclamation of secondary structure¹⁰, potentially reducing the aggregation. Therefore, the formation of C-BSA aggregates upon cationization of BSA should not hinder the formation of a solvent-free liquid BSA.

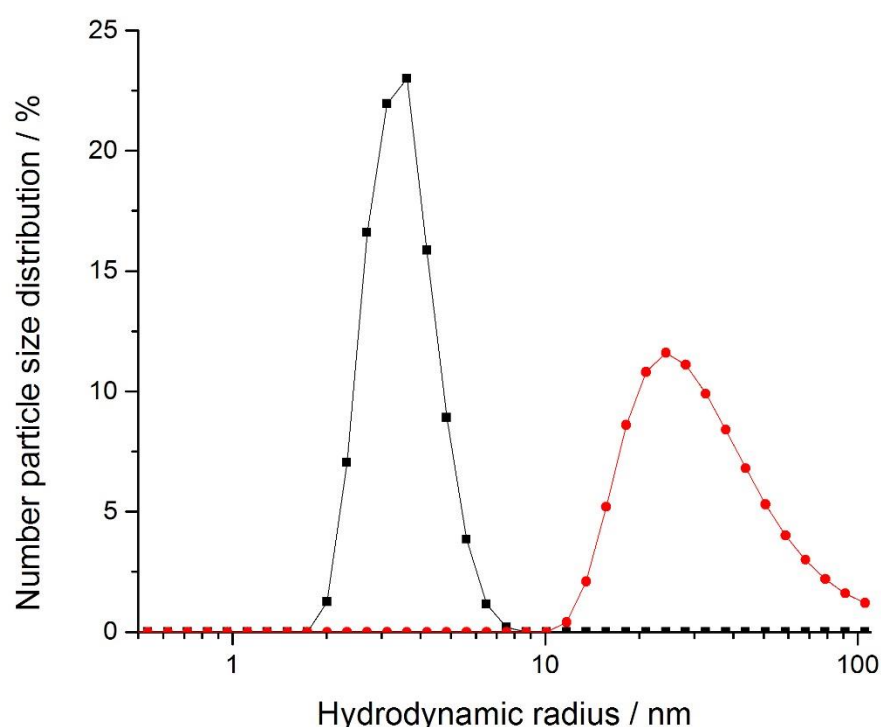


Figure 3.4: The number particle size distribution for aqueous BSA (black) and aqueous cationized BSA (red). [Conditions: unbuffered solutions at pH values of 6-6.5, with proteins concentrations of 0.5 mg.mL⁻¹, 25 °C]. Data displayed is an accumulation of at least 10 scans.

3.2.1.3.2 DLS results for PK and C-PK

The calculated results for the hydrodynamic radii for PK and C-PK were 2.3 nm and 3.1 nm (Figure 3.5) respectively, with the radius of PK fitting with previous literature values¹¹. This showed that the cationization of PK increased the hydrodynamic radius by less than 1 nm relative to PK. In contrast to BSA, the cationization of PK has only slightly increased the hydrodynamic radius, with the increase being attributed to the length of the additionally attached amines and potentially a small amount of unfolding of the secondary structure. The lack of larger structures in the sample (<1% above 10 nm) signified that PK did not aggregate upon cationization, indicating that the high surface charge present and any potential unfolding that occurred did not initiate aggregation.

The lack of C-PK aggregation supports the hypothesis that it was the partial unfolding of BSA that led to aggregation of C-BSA rather than the presence of highly charged proteins. This may demonstrate that the cationization of PK forms a more stable structure than that of C-BSA, however, a structural study was necessary to determine this.

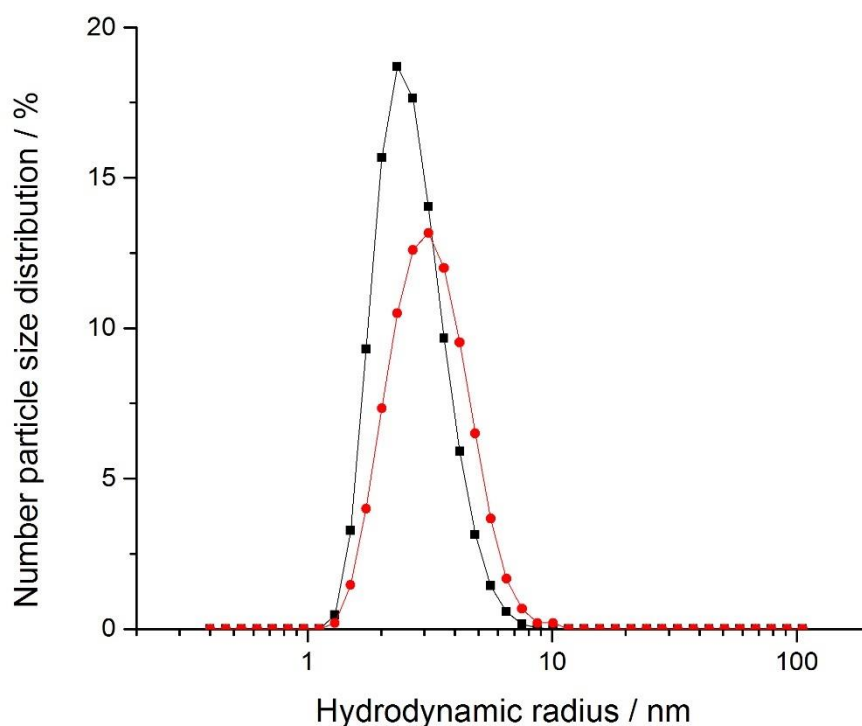


Figure 3.5: The number particle size distribution for aqueous PK (black) and aqueous cationized PK (red). [Conditions: unbuffered solutions at pH values of 6-6.5, with proteins concentrations of 0.5 mg.mL⁻¹, 25 °C]. Data displayed is an accumulation of at least 10 scans.

3.2.1.4 Secondary structure analysis through circular dichroism spectroscopy (CD)

As previously detailed, CD spectroscopy is a technique that can be utilised to study the secondary structure of a protein, with the spectra being deconvoluted in order to estimate the percentages of secondary structure elements. The technique is used here to study the effect of cationization on the secondary structure of BSA and PK relative to the native species of each respective protein.

3.2.1.4.1 Structure analysis of BSA and C-BSA

The structure of BSA is known to be dominated by α -helical protein moieties, having characteristic minima at 222 nm and 208 nm along with a characteristic maximum at 195 nm. It is evident from the spectra obtained that the cationization of BSA to form C-BSA causes an alteration in secondary structure relative to native-BSA (Figure 3.6). This is seen as a reduction in intensity in all three characteristic peaks along with a blue shift in the parallel $\pi \rightarrow \pi^*$ transition (208 to 205 nm) and a blue

shift in the $\pi \rightarrow \pi^*$ transition (193 to 190 nm). The decrease in intensity is not uniform across the characteristic peaks, with a significant reduction in the ellipticity of the negative 222 nm peak relative to the negative 205 nm peak, altering the peaks ratios. This leads to a $\theta_{222}/\theta_{205}$ peak ratio decrease, from ≈ 1 to 0.69, an indication of the α -helices becoming unstable and forming 3_{10} helices¹², although further structure studies would be required to confirm this. The large decrease in intensity of the ≈ 195 nm transition relative to the other two peaks could be indicating a decrease in the positive perpendicular $\pi \rightarrow \pi^*$ transition peak (related to α -helical moieties) and an increase in the negative $\pi \rightarrow \pi^*$ transition (related to random coil moieties). Such a peak shift could be a further indication of the loss of α -helical structure and an increase in random coils.

It is clear from the deconvoluted CD data (Table 3.3) that cationization of native-BSA has a negative impact on the secondary structure; α -helical content decreases from 37% to 18%, with a concurrent increase in the unordered and β -sheet content, from 37% to 42% and from 14% to 26% respectively. The turn content does not alter to a large degree, indicating that the majority of α -helical content transitioned to become β -sheets or unordered structures. A decrease in CD signal intensity of this magnitude, accompanied with the estimated alterations to the secondary structure, would allude to a large unfolding event occurring once the BSA is cationized.

Deviation in the secondary structure relative to BSA to this degree occurs due to the alteration of surface charge caused by the covalent addition of DMAPA to 100% of the negatively charged external residues. The replacement of all negative surface charges with positive charges would disrupt all of the structurally significant ion pairings and the 8 known salt bridges that are solvent accessible³. All of these previously structurally stabilizing effects would transform to become either non-stabilising or destabilizing, with the now numerous same-charge electrostatic interactions adding additional repulsive forces to the delicate force balance of a folded protein. The additional repulsion leads to conformational shifts as the protein rearranges into the new lowest energy conformation. This would lead to the exposure of numerous (originally internalized) residues that display groups that are predominantly hydrophobic.

The exposure of these hydrophobic groups to an aqueous environment often leads to aggregation due to the positive shielding afforded by the hydrophobic regions of proximal proteins interacting, it being energetically favourable to do so over interacting with water molecules. These interactions can often take the form of inter-protein β -sheet formations at the expense of α -helical conformations, which in turn can lead to a further increase in aggregation, creating a positively reinforced feedback mechanism^{13,14,15}. The evidence for this occurring with CBSA is in the increase in β -sheet content and the decrease in α -helical content (Table 3.3), and even though the changes are not equal in magnitude it can be postulated that there is conversion of α -helices to β -sheets. A similar effect is seen in amyloid fibril formation, whereby proteins that are insoluble, often due to incorrect folding or translation errors, interact through hydrophobic interactions and form inter-protein β -sheets, eventually inducing the formation of fibrils that lead to a number of different diseases^{16,17}.

It is, therefore, likely that C-BSA, being highly unfolded relative to BSA, forms aggregates, a theory that is supported by the DLS data, with C-BSA having a larger average hydrodynamic radius and evidence of numerous larger aggregates in the system. This could potentially have caused issue in the formation of a solvent-free liquid BSA due to the reasons detailed earlier, however, as discussed earlier, this was not the case.

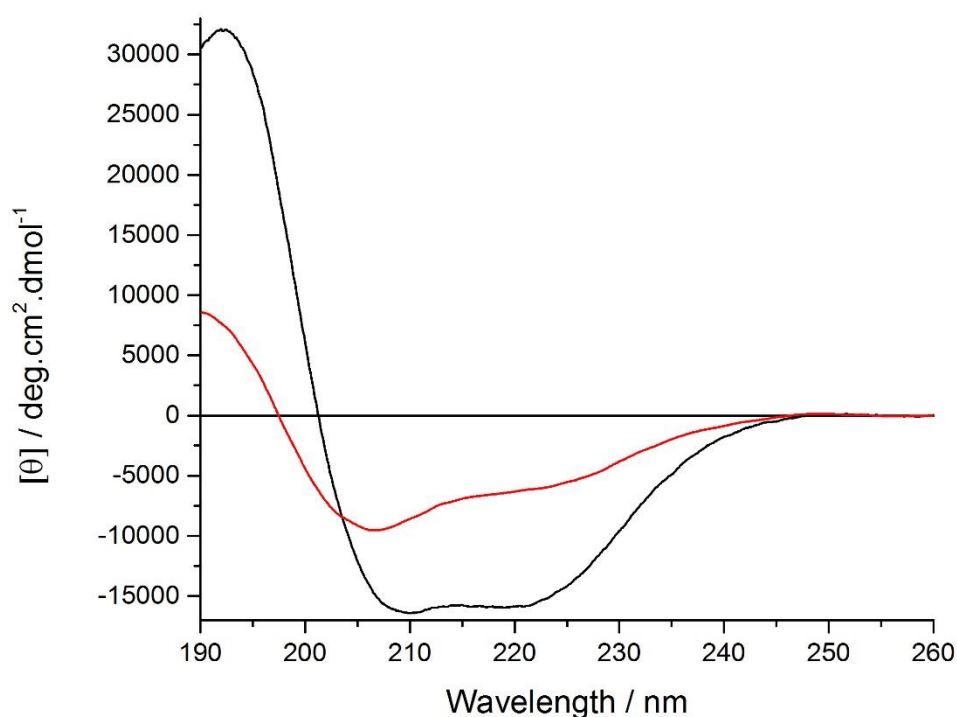


Figure 3.6: Mean residue ellipticity against wavelength showing far-UV CD spectra of aqueous BSA (black) and aqueous cationized BSA (red) at 25 °C between 190-260nm.

Table 3.3: Table of secondary structure content calculated from CD spectra displayed in Figure 3.6.

Sample	α -Helix / %	β -Sheet / %	Turns / %	Other/ %
Aqueous C-BSA	18	26	14	42
Aqueous native-BSA	37	14	12	37

3.2.1.4.2 Structure analysis of PK and C-PK

The secondary structure of proteinase K is known to contain an extended β -sheet that is decorated by a number of α -helices, with the CD spectra displaying broad characteristic minima at 222 nm and 210 nm along with a characteristic maximum at 195 nm. Once cationized, the intensity of all three characteristic peaks unexpectedly increased, with a narrowing of the negative $n \rightarrow \pi^*$ transition peak and a blue shift of the parallel $\pi \rightarrow \pi^*$ transition peak (210-208 nm) and of the $\pi \rightarrow \pi^*$ transition peak (195-190 nm) (Figure 3.7). An increase in all characteristic peaks indicates a potential increase in α -helical and/or β -Sheet moieties and a reduction in random coils. Similar to the cationization of BSA, the change in intensity is not uniform across the characteristic peaks, with a significant increase in the ellipticity of the negative 210 nm peak relative to the negative 222 nm peak, altering the peaks ratios. This gives a $\theta_{222}/\theta_{208}$ ratio increase (1.2 to 1.5), indicating an increase in α -helices and also the potential formation of coiled coils within the C-PK structure^{18,19}, although further structure studies would be required to confirm this.

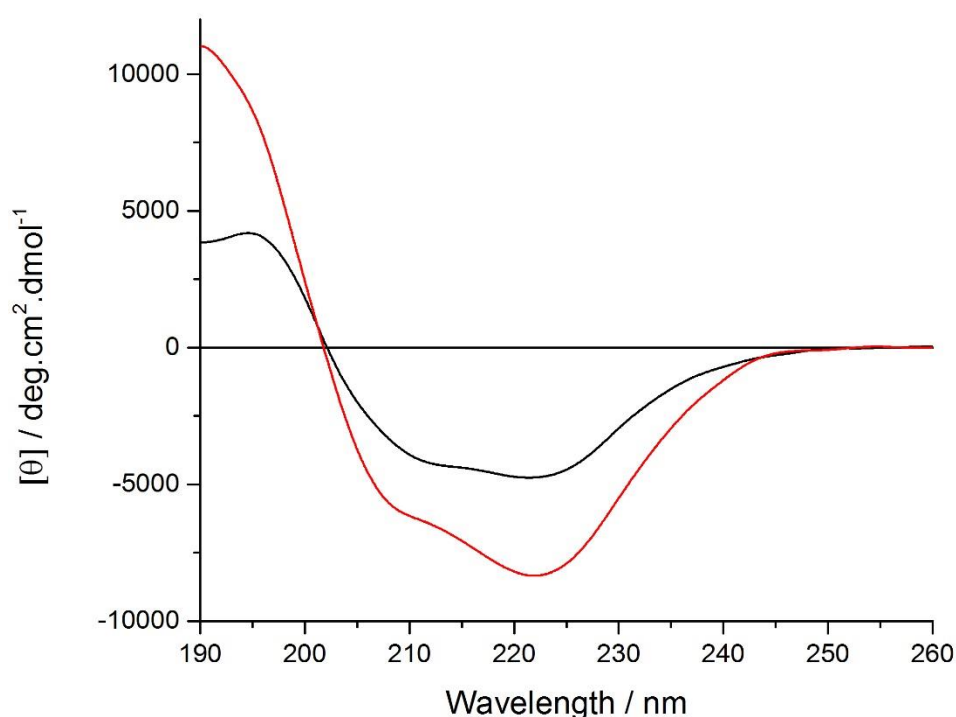


Figure 3.7: Mean residue ellipticity against wavelength showing far-UV CD spectra for aqueous PK (black) and aqueous cationized PK (red) at 25 °C between 190-260nm.

Table 3.4: Table of secondary structure content calculated from CD spectra displayed in Figure 3.7.

Sample	α -Helix / %	β -Sheet / %	Turns / %	Other/ %
Aqueous C-PK	18	33	13	37
Aqueous native-PK	10	34	13	43

It is also clear from the deconvoluted CD data (Table 3.4) that the surface cationization increases α -helical content (10-18%) with a concurrent drop in unordered secondary structure content. Interestingly, this indicates that stabilization occurs in the structure once the protein is highly cationized. The turn content does not alter indicating that the majority of additional α -helical content transitioned from unordered structures, further proof of the increased structural stability of the cationized PK relative to native-PK. As the deconvoluted data is an estimation of secondary structure content it is more likely that the increase in secondary structure is spread across both α -helical and β -sheet moieties.

An increase in the secondary structure of PK was unexpected. All previous work relating to the formation of solvent-free liquid proteins have concluded that the cationization of the protein leads to a drop in secondary structure due to the loss of surface ion pairings and salt bridges. However, as is seen in numerous works on site-specific residue mutagenesis and site-specific charge changes, replacing negative surface charges with positive ones can lead to more stable structures²⁰. Although in the majority of these studies only 1-2 sites are altered, the theory of why the stabilization occurs is still relevant. The charges present on the surface, whilst generally aiding in stabilization through opposite-charge attractive interactions, can be destabilizing. Therefore, the removal of these charges, and even a conversion of their charge, can lead to an overall stabilization effect.

Any deviation from the native secondary or tertiary structure of a protease has great potential to alter the enzyme activity due to conformational changes altering the protease/substrate interaction mechanism. As well as conformational changes, the deviation from native charge density, especially proximal to the active site and catalytic binding cleft, will alter how the protease and substrate interact. As is evident from the observed increase in CD signal intensity and from the calculated deconvoluted data, there is an increase in secondary structure of PK induced by cationization. It is also clear that there is an increase in global charge density, seen through an increase in zeta potential. These two alterations from native-PK could potentially alter the activity of PK towards the substrate (BSA), with all activity related discussion in Chapter 4. However, the residues forming the substrate binding cleft as well as the residues making up the catalytic triad within a protease are known to be flexible²¹. This appreciable flexibility in the key interaction sites may permit the protease to be cationized, with the subsequent structure and charge changes, whilst retaining some level of enzyme activity. It is interesting to note that the cationization of PK did not initiate aggregation of the highly charged protein as was seen in CBSA. The reason for this is linked to the increase in α -helical content, a trait known to inhibit protein aggregation¹³. The increase in positive surface charge, along with a commensurate decrease in negative surface charge stabilised the secondary structure of the protein, the opposite effect to that seen with the cationization of BSA. This might be an effect of the cationization efficiency not being 100%, with a number of negatively charged site remaining unmodified. There is also no increase in inter- or intra-protein β -sheet content, another indicator that no aggregation has occurred. Whilst CPK has a higher level of secondary structure, with fewer

unordered areas of the protein, this does not necessarily correlate to an increase in protease activity and could instead negatively impact the activity.

3.2.1.5 Thermal denaturation

Proteins are known to lose or alter their secondary structure due to environmental changes such as temperature and pressure, with each protein denaturing at a protein specific temperature or pressure. The denaturation dynamics is unique to each protein. Accordingly, a cationized protein may have different unfolding thermodynamics the relative native protein. One method for quantifying the thermal stability of a protein is to calculate the half denaturation temperature (T_m): being the temperature at which both the folded and unfolded states are equally populated, at which half of the population is 100% folded and the other half is 100% unfolded, or at some point between the two. The calculation of T_m relies on the assumption of cooperative protein unfolding, whereby a protein rapidly transitions from being in a folded state to being in an unfolded state. To calculate the T_m , and subsequently the enthalpy (ΔH_m) and entropy (ΔS_m) of half denaturation, the characteristic 222 nm CD peak obtained from a set of temperature-dependent CD spectra is used as an order parameter (see section 2.1.4) for the fraction of denatured protein as a function of temperature.

Equilibrium thermal denaturation experiments were conducted allowing for quantitative thermodynamic comparison between the cationized and native species of the proteins. Alongside these, the deconvolution of thermal CD data allows for alterations to the secondary structure motifs to be tracked, the percentages of which can be plotted against temperature, providing more unfolding data with which to compare the native and cationized proteins.

3.2.1.5.1 Thermal denaturation of BSA and C-BSA

The CD thermal denaturation curve for native-BSA in Figure 3.8B has a clear sigmodal shape, indicating that the unfolding of BSA due to temperature increase was a two-stage mechanism. The T_m for BSA was calculated to be 72 °C. This calculated value for T_m corresponds to a crossover of dominant secondary structure motif, from α -helices to β -sheets at ≈ 70 °C (Figure 3.8D). Due to the experimental limitations, BSA being aqueously solvated, BSA was only able to be shown to retain secondary structure up to 95 °C, however, it is known that BSA can retain a level of secondary structure up to temperature as high as 130 °C. In this case it is clear from denaturation curve and the deconvoluted thermal CD spectra that there was little change in the secondary structure up to ≈ 60 °C, with only a marginal decrease in 222 nm signal intensity and a small percentage of α -helices converting to β -sheets. This indicates that BSA retains native secondary structure up to 60 °C and is thermally resistant to temperature increase up to this temperature. As the temperature was increased past 60 °C there was a rapid alteration in secondary structure, with a large decrease in 222 nm signal intensity and in the percentage of α -helices with a commensurate increase in β -sheet percentage. The large decrease in 222 nm signal is also mirrored in both the 195 and 208 nm peaks (Figure 3.8A). After ≈ 80 °C the rate

of secondary structure change decreases dramatically, this signifying that BSA has entered an unfolded state, with any further increase in temperature having minimal effect on the remaining secondary structure. It is likely that the unfolding of α -helices with BSA due to heating past 60 °C could drive the formation of inter-protein β -sheet formation, as seen in the cationization of BSA. This could dramatically speed up the unfolding of BSA, altering the unfolding dynamics relative to a single BSA unit unfolding without other BSA units surrounding it. The formation of these β -sheets could lead to aggregation, also altering the unfolding dynamics.

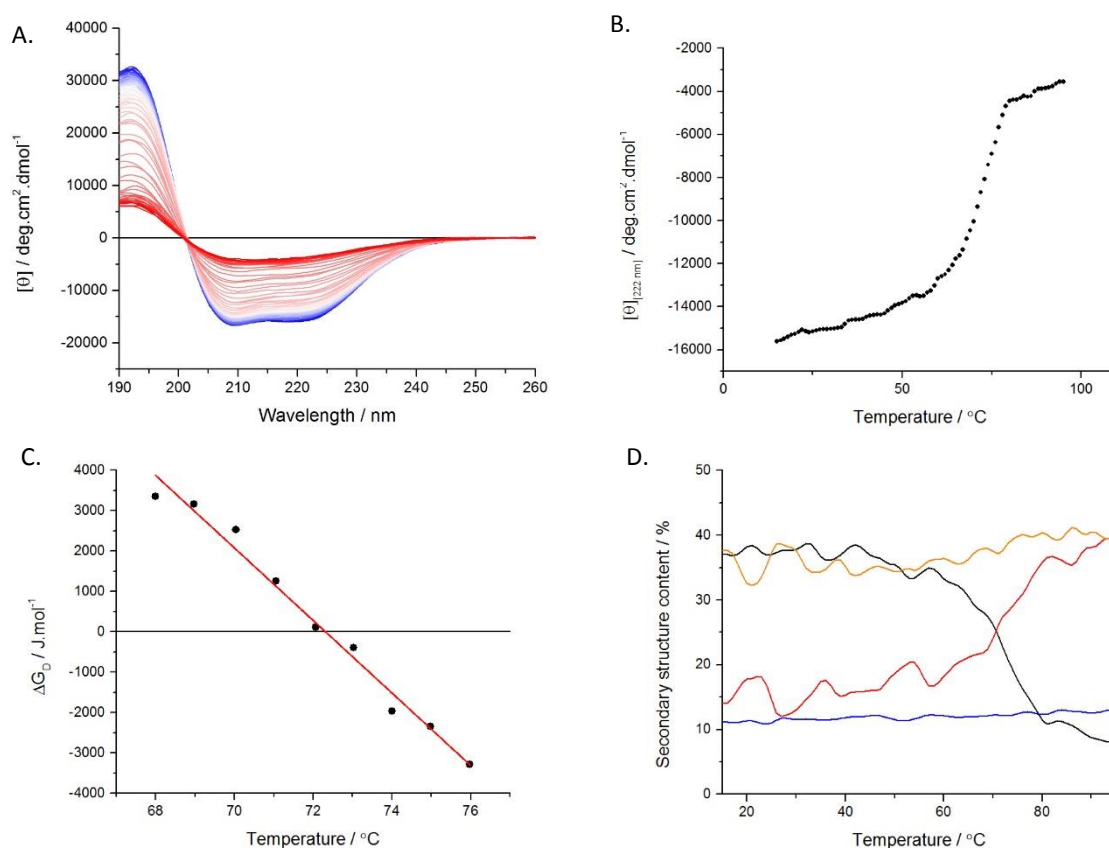


Figure 3.8: **A.** Mean residue ellipticity against wavelength showing far-UV CD spectra for aqueous BSA, with the gradient from blue to red indicating an increase in temperature from 20 – 95 °C, **B.** Mean residue ellipticity at 222 nm against temperature (20 – 95 °C) for aqueous BSA, **C.** plot of free energy of denaturation (ΔG_D) against temperature displaying the transition region for BSA (the linear fit (red) was used to calculate the thermodynamic parameters), and **D.** secondary structure percentage change against temperature (20 – 95 °C) for aqueous BSA (secondary structure percentage calculation detailed in Chapter 2), with the percentage of α -helices (black), β -sheets (red), random coil (orange), and turns (blue).

Table 3.5: Table displaying the transition melting temperature, enthalpy change for melting transition, and entropy change for melting transition (T_m , ΔH_m , and ΔS_m respectively) for aqueous BSA.

Sample	T_m / °C	ΔH_m / kJ mol ⁻¹	ΔS_m / J K ⁻¹ mol ⁻¹
Aqueous native-BSA	72 ± 3	320 ± 14	930 ± 9

In contrast to BSA, C-BSA did not show a clear transition in secondary structure percentage nor in the temperature dependent CD, indicating a non-cooperative unfolding mechanism that can be attributed to a stabilized molten globule-like conformational state (Figure 3.9A and B)¹⁰. Molten globule protein states can be induced by an increase in surface charge repulsion and the disruption of salt bridges, both of which occur during the cationization stage of the synthesis. This disruption leads to a loosely bound tertiary structure that can often retain a large percentage of the secondary structure relative to the native. The thermodynamic unfolding pathway of a molten globule protein differs from that of a native protein as the more loosely bound peptide chains can dissipate energy through additional vibrations. This has the effect of potentially increasing the T_m of the protein. With BSA displaying a T_m of 72 °C, with an increase could take it above the solvent temperature limit. The formation of CBSA aggregates, shown to occur during the formation of CBSA, could also alter the unfolding dynamics of CBSA relative to BSA as the protein is known to already be in a highly unfolded state.

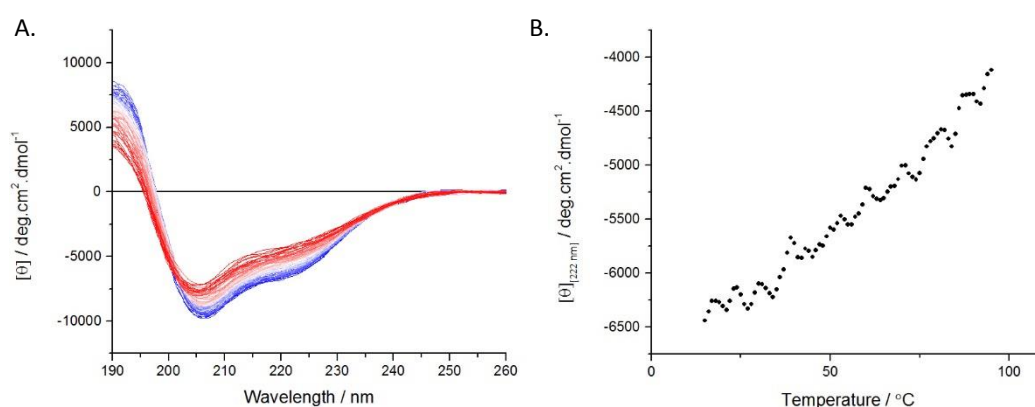


Figure 3.9: **A.** Mean residue ellipticity against wavelength showing far-UV CD spectra for aqueous C-BSA, with the gradient from blue to red indicating an increase in temperature from 20 – 95 °C. **B.** Mean residue ellipticity at 222 nm against temperature (20 – 95 °C) for aqueous C-BSA.

As C-BSA did not cooperatively unfold, the calculation model for T_m is not specifically applicable and doesn't yield accurate thermodynamic values. Previous work conducted on cationized proteins detailed a set of chemical denaturation experiments that allow for an alternative method for the calculation of the T_m of C-BSA¹⁰. Chemical denaturation has the effect of disrupting local electrostatic interactions, much in the same way as increasing the temperature of the solution does. Therefore, chemical denaturation can be used to simulate an increase in the temperature the solution, allowing for investigation past the temperature limit of the solvent to access the fully-denatured state.

The chemical denaturation of C-BSA was performed using the denaturant urea, a denaturant often used in high concentrations (8M) to fully denature a protein. Here, a range of lower urea concentrations was used on C-BSA, with the chemically denatured sample then being subjected to a thermal CD scan between 20 and 95°C and 190-260nm. If a sigmoidal curve was present in the denaturation curve, the T_m for that concentration can be calculated as detailed later. By plotting the calculated T_m against the concentration of urea and extrapolating back to a urea concentration of 0

M, the T_m for non-chemically denatured C-BSA could be approximated. This approximation is only correct if the increase in urea concentration increases the level of denaturation linearly.

Chemical denaturation of C-BSA between urea concentrations of 1.5-2.5 M recovered the sigmoidal nature of the denaturation curve (example in Figure 3.10), indicating that the unfolding is in line with the two-stage protein unfolding model required for the calculation of T_m . Urea concentrations higher than 2.5 M did not display a sigmoidal denaturation curve, indicating that the high concentration of urea had unfolded the protein into a highly unfolded state, with any further increase in temperature having little effect on the structure of CBSA. Further evidence to this is the decrease in 222 nm signal intensity as the concentration of urea increased, plateauing after a urea concentration of 3 M, implying that concentrations higher than 3 M the protein is in a fully-unfolded state. As the calculated T_m for various concentrations of urea increased linearly (Figure 3.11) the data was then extrapolated to yield a T_m of 89 ± 1 °C. This approximated T_m is 17 °C higher than that of BSA, giving evidence towards the theory that the highly cationized protein enters a loosely bound molten globule state that has a higher resistance to thermal degradation. This higher thermal resistance comes from the protein already being in a low entropic state (more unfolded relative to BSA), with the increase in temperature therefore not having as large an impact on the state of the conformation. Higher temperature environments drive the high entropic state of a fully folded protein towards the lower entropic state of being unfolded. Therefore, if the protein begins in a lower entropic state it takes more energy to alter the conformation to a large degree.

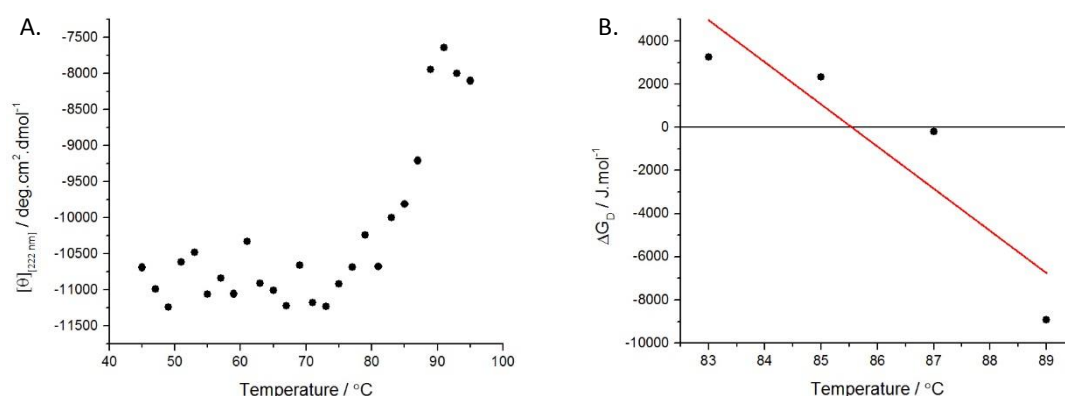


Figure 3.10: **A.** Mean residue ellipticity at 222 nm against temperature (20 – 95 °C) for C-BSA in 2 M urea and **B.** Plot of free energy of denaturation (ΔG_D) against temperature displaying the transition region for C-BSA in 2 M urea. The linear fit (red) was used to calculate the thermodynamic parameters.

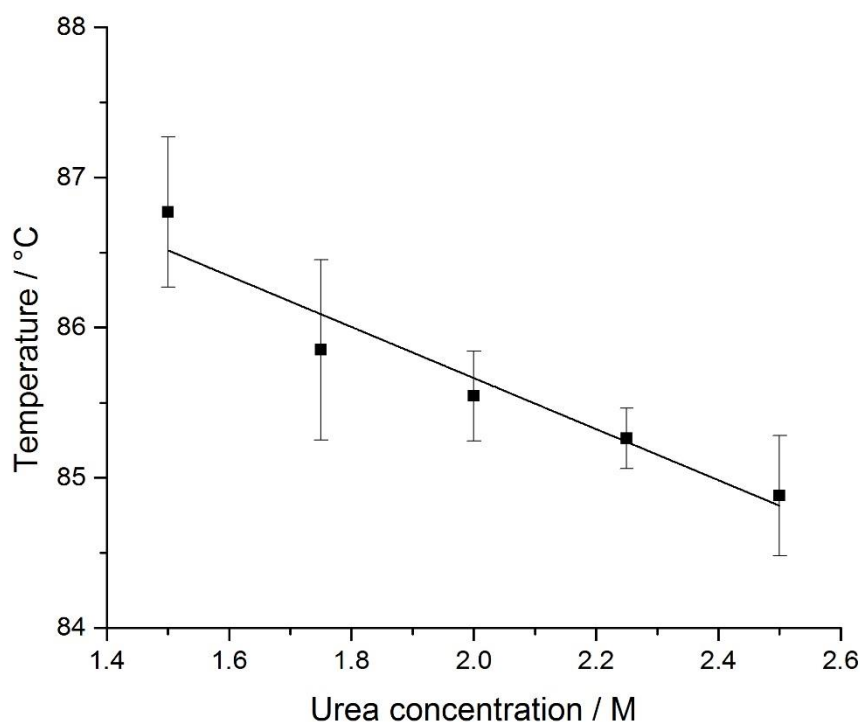


Figure 3.11: Melting transition temperature against urea concentration for various C-BSA samples denatured with a range of urea concentrations.

3.2.1.5.2 Thermal denaturation of PK and C-PK

The circular dichroism thermal denaturation experiment for native-PK yielded markedly different results to that of native-BSA, with the structure rapidly degrading from a highly ordered secondary structure protein to a more random coil dominated one, retaining no secondary structure motifs past 60 °C (Figure 3.12A). This loss of secondary structure motifs is not evident in the deconvoluted CD data, instead this data showed only small changes across the temperature range. As the full wavelength and temperature CD spectra clearly displayed a drastic shift in characteristic spectra, the values of the deconvoluted data were disregarded and merely the trends are discussed. The thermal denaturation 222 nm curve (Figure 3.12B) displayed a double sigmoidal characteristic with a plateaued region between the two sigmoids, in contrast with the single sigmoid displayed in native-BSA. This double sigmoid indicates a three-stage cooperative unfolding mechanism with the presence of a stable equilibrium intermediate state; transferring from a native state (N) to an intermediate state (I) and subsequently on to a denatured state (D). The intermediate state is created due to the increased thermal energy of the system unfolding PK, with PK reordering its secondary structure and finding a new temporarily-stable global energy minimum at a specific temperature (≈ 40 °C). As the thermal energy was increased further the PK again begins to unfold, this time continually unfolding until it reached a fully denatured state with a CD spectrum that depicted a highly random coil dominant protein.

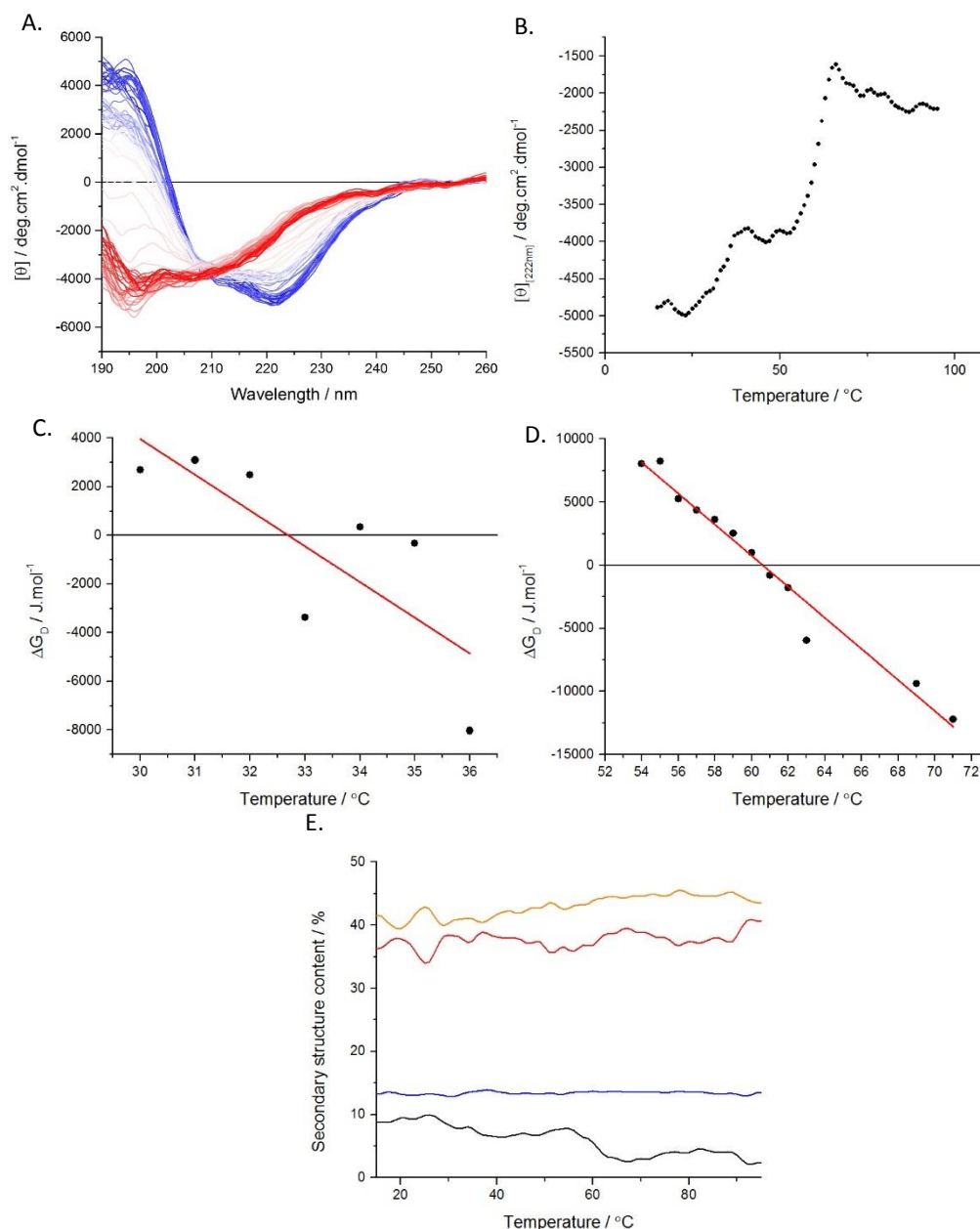


Figure 3.12: **A.** Mean residue ellipticity against wavelength showing far-UV CD spectra for aqueous PK, with the gradient from blue to red displaying an increase in temperature from 20 – 95 °C, **B.** Mean residue ellipticity at 222 nm against temperature (20 – 95 °C) for aqueous PK, **C.** plot of free energy of denaturation (ΔG_D) against temperature displaying the transition region for PK_a (the linear fit (red) was used to calculate the thermodynamic parameters), **D.** plot of free energy of denaturation (ΔG_D) against temperature displaying the transition region for PK_b (the linear fit (red) was used to calculate the thermodynamic parameters), and **E.** secondary structure percentage change against temperature (20 – 95 °C) for aqueous PK (secondary structure percentage calculation detailed in Chapter 2), with the percentage of α -helices (black), β -sheets (red), random coil (orange), and turns (blue).

Due to the presence of two sigmoidal features in the thermal denaturation 222 nm graph (Figure 3.12B), the calculation for T_m was conducted for each of the two transition curves, yielding two separate values for T_m along with two sets of thermodynamic values (one for each transition) being labelled T_{ma} and T_{mb} for the first and second transition respectively. The first transition (T_{ma}) occurs at 33 °C (Table 3.6), with the intermediate state stabilizing at 40 °C and remaining so until 55 °C. Unlike

BSA, the secondary structure content (Figure 3.8D) did not show a large decrease in either α -helices or β -sheets at this temperature despite the observed drop in 222 nm signal intensity. This could potentially indicate that the structure of the intermediate state is native-like although possibly less densely packed than PK at 25 °C. This is supported by the 190-260nm scan of the sample at 40 °C which is similar in curvature, although not in signal intensity, to one at 25 °C. T_{m2} occurs at 61 °C (Table 3.6), with a drastic decrease in 222 nm CD signal, indicating a rapid change in secondary structure which is observed by a large change in spectra curvature in the 190-260nm wavelength scan. This secondary transition is evident in the secondary structure content, with a slight decrease in α -helices and an increase in β -sheets and random coils. However, the observed decrease in secondary structure motifs is not as large as would be expected from the large change in CD intensity displayed in the thermal denaturation spectrum (Figure 3.12D). Additional, more complex, structural analysis would be required to obtain a greater degree of accuracy regarding the secondary structure, as the deconvolution estimate of this data did not yield a representative set of secondary structure percentages.

Regarding the entropy and enthalpy of transitions, it would seem apparent that the native state was more stable in relation to the intermediate species than the intermediate species was towards the denatured species, evident in the higher entropy and enthalpy of transition for T_{ma} compared to T_{mb} (Table 3.6). This implies that at the increased temperature the re-ordered configuration attained formed a less stable construct in relation to the native state, with a lower entropy change upon transition. This result was expected, however, the high retention of stability displayed in the magnitude of transitional entropy and enthalpy of the intermediate species was unexpected. This indicates that the intermediate species is native-like in thermal stability as well as the structure and may possibly help retain the activity of PK at elevated temperatures, although temperature-dependant activity studies would be required to confirm this hypothesis.

Table 3.6: Table displaying the transition melting temperature, enthalpy change for melting transition, and entropy change for melting transition (T_m , ΔH_m , and ΔS_m respectively) for aqueous PK

Sample	$T_m / ^\circ\text{C}$	$\Delta H_m / \text{kJ mol}^{-1}$	$\Delta S_m / \text{J K}^{-1} \text{mol}^{-1}$
Aqueous native-PK b	61 ± 6	410 ± 29	1200 ± 87
Aqueous native-PK a	33 ± 16	450 ± 160	1500 ± 500

Cationization of PK, as previously detailed, appeared to increase the α -helical content of PK, indicating that the secondary structure is more stable in the cationized form than in the native form at 25 °C. This structure alteration could potentially increase the thermal stability of PK, with the more favourable interactions present within the structure resisting thermal alteration to a higher temperature. However, as seen with C-BSA, there is also the potential to form a molten globular protein, which is likely to have increased thermal stability, potentially masking other causes for an increase in thermal stability.

It is clear from the C-PK thermal denaturation 222 nm CD spectrum (Figure 3.13B) that cationization altered the thermal unfolding dynamics of native-PK. It is known that, although the double sigmoidal characteristic was retained, a molten globular-like protein was not created due to the sharp transitions observed in the curve. T_{ma} for C-PK occurs at 39 °C, this was 6 °C higher than that of native-PK. This increase in initial thermal stability indicates that the configuration of the C-PK is initially more thermally stable than that of native-PK, the structure resisting alteration until a higher temperature. This trend would appear to continue with an intermediate state being stabilizing 46 °C, this again being 6 °C higher than the intermediate of native-PK. As the 222 nm intensity of the C-PK stabilized intermediate state is higher than that of the stabilized PK intermediate state ($\approx 4500 \text{ deg.cm}^2.\text{dmol}^{-1}$ compared to $\approx 4000 \text{ deg.cm}^2.\text{dmol}^{-1}$), it is logical to surmise that the intermediates from each system are of different configurations.

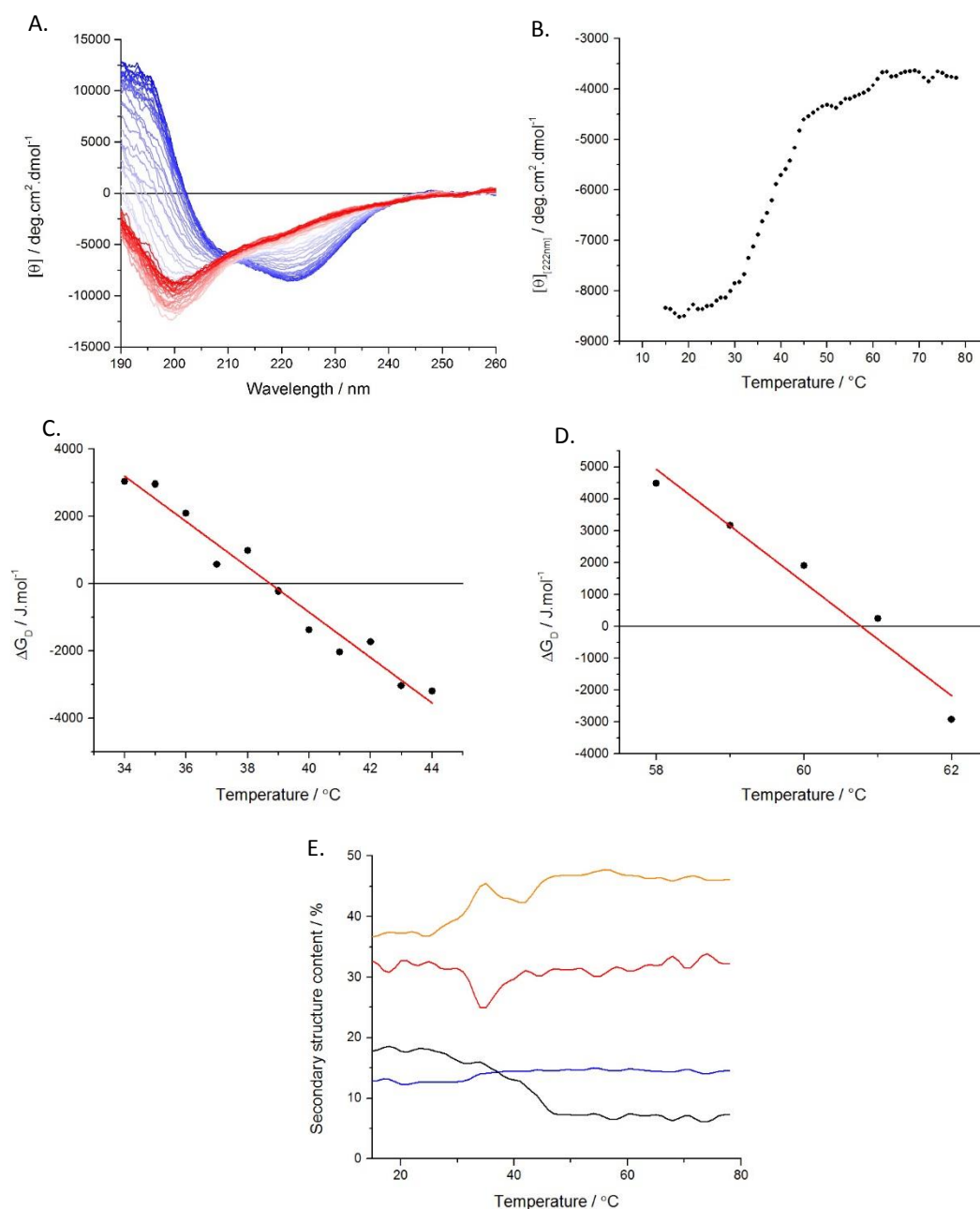


Figure 3.13: **A.** Mean residue ellipticity against wavelength showing far-UV CD spectra for aqueous C-PK, with the gradient from blue to red displaying an increase in temperature from 20 – 95 °C, **B.** Mean residue ellipticity at 222 nm against temperature (20 – 95 °C) for aqueous C-PK, **C.** plot of free energy of denaturation (ΔG_D) against temperature displaying the transition region for C-PK_a (the linear fit (red) was used to calculate the thermodynamic parameters), **D.** plot of free energy of denaturation (ΔG_D) against temperature displaying the transition region for C-PK_b (the linear fit (red) was used to calculate the thermodynamic parameters), and **E.** secondary structure percentage change against temperature (20 – 95°C) for aqueous C-PK (secondary structure percentage calculation detailed in Chapter 2), with the percentage of α -helices (black), β -sheets (red), random coil (orange), and turns (blue).

It is clear from the thermal denaturation curve that the T_{ma} transition for C-PK is the more dominant denaturation with a 50% decrease in signal intensity. Whereas for native-PK the dominant transition was T_{mb} , again having a 50% decrease in intensity. Unlike PK, T_{ma} is also evident in the deconvoluted data of C-PK, with a clear decrease in secondary structure content at around 39 °C; specifically, a decrease in α -helices and β -sheets and an increase in the percentage of random coils. Whilst the T_{ma} of both the PK and C-PK occur at a similar temperature the result of this initial unfolding is markedly

different. PK unfolds into a relatively native-like intermediate configuration, whereas C-PK unfolds into an almost fully denatured intermediate configuration. The C-PK intermediate state also gradually loses structure as the temperature increased, a trait in contrast to the PK intermediate state, indicating that the PK intermediate state is more thermally stable than that of C-PK. Despite the differences in starting configuration, initial unfolding temperature, and intermediate state, the T_{mb} of both native-PK and C-PK is the same, being 61 °C. The gradual loss of structure in the C-PK intermediate considerably increased the error in this calculation, meaning the calculated T_{mb} s of PK and C-PK might not actually be as close as it appears.

Regarding the entropy and enthalpy of transitions, it would seem apparent that the native state was less stable in relation to the intermediate state than the intermediate state was towards the denatured state; this due to the lower entropy and enthalpy of transition for T_{ma} compared to T_{mb} (Table 3.7). With ΔS_{ma} being lower than ΔS_{mb} this means that there was a smaller change in entropy to enter the intermediate state from the native state than there was for the intermediate state to enter the denatured state. This implies that the configuration attained through the re-ordering at increased temperature formed a more stable construct in relation to the native state, having a much larger entropy change and a higher release of energy when transitioned. The ΔH_{ma} and ΔS_{ma} for C-PK are half that of the PK species, meaning the PK native state is more stable towards its intermediate species. More interestingly, for PK, the ΔH_{mb} and ΔS_{mb} for T_{mb} are 0.33 times lower, meaning the C-PK intermediate state is more stable in relation to the C-PK denatured state than the PK intermediate is in relation to the PK denatured state. All of the evidence presented indicates that the addition of multiple charges to the surface alters the thermodynamics of PK. Initially displaying an increased stability at 25 °C through increased percentages of secondary structure motifs and an increased T_{ma} , the intermediate species formed is less stable to thermal alterations than that of PK, with the appearance of some retained native-like core thermodynamics.

Table 3.7: Table displaying the transition melting temperature, enthalpy change for melting transition, and entropy change for melting transition (T_m , ΔH_m , and ΔS_m respectively) for aqueous PK and C-PK.

Sample	$T_m / ^\circ\text{C}$	$\Delta H_m / \text{kJ mol}^{-1}$	$\Delta S_m / \text{J K}^{-1} \text{mol}^{-1}$
Aqueous C-PK b	61 ± 10	590 ± 73	1800 ± 220
Aqueous C-PK a	39 ± 4	220 ± 15	700 ± 49
Aqueous native-PK b	61 ± 6	410 ± 29	1200 ± 87
Aqueous native-PK a	33 ± 16	450 ± 160	1500 ± 500

3.2.1.6 Conclusion for cationization

Cationization of BSA to form C-BSA and of PK to form C-PK through the covalent addition of amides to the surface of the proteins had a similar effect to both proteins regarding the surface charge alteration but drastically different impacts to secondary structure and thermal stability of each protein. It is evident from mass spectrometry (MS) measurements that the cationization added a detectable

number of DMAPA molecules to the surface of BSA and PK, being 100% and 89% efficient respectively. This was supported by zeta potential (ZP) measurements displaying a charge density increase, *i.e.* towards being more positive, in the case of both proteins. Once cationized, C-BSA displayed the formation of aggregate structures with large hydrodynamic radii thought to manifest due to the unfolding of BSA exposing the hydrophobic core residues. In contrast, C-PK did not form aggregates, with the average hydrodynamic radii only marginally increased with respect to PK. Drastically different structures are formed by increasing the net surface charge (cationization) of the native proteins BSA and PK caused by the disruption of surface ion pairings and salt bridges through conversion of the negative sites to positive.

Circular dichroism spectrometry (CD) of C-BSA at 25 °C displayed a drastic decrease in signal intensity relative to BSA, along with a decrease in secondary structure present from the deconvoluted CD data. This indicated a large degree of protein unfolding due to cationization, evidence being supported by dynamic light scattering (DLS) measurements showing a larger average particle radius and evidence of large aggregates, thought to be generated by hydrophobic interactions between exposed core BSA residues forming inter-protein β -sheets. Thermal CD spectroscopy displayed the formation of a molten globular-like protein that had a 17 °C higher melting transition temperature than BSA due to the lower-density packing conformation the protein took, further evidence supporting the unfolding of BSA. This indicates that the increase in positive surface charge and subsequent alteration to the electrostatic stabilising forces present on the surface of C-BSA had a detrimental impact to the structure of BSA and also altered the thermodynamic properties of the protein.

In contrast to C-BSA, the 25 °C CD of C-PK displayed an increase in signal intensity relative to PK, along with an increase in secondary structure present from the deconvoluted CD data, namely an 8% increase in α -helical content. This indicated that cationization could increase the packing in the case of C-PK, with the potential of the removal of β -sheet stabilising salt-bridges during cationization, allowing the formation of α -helices. Unlike the aggregation with C-BSA, DLS of C-PK did not display evidence of aggregation. Again, in contrast to C-BSA, thermal CD spectroscopy did not display the formation of a molten globular-like protein. Rather, the change to the thermodynamics observed was due to an alteration to the stability of a partially folded intermediate species, indicating that a different intermediate species was formed in the thermal degradation pathway for C-PK relative to PK. The entropy and enthalpies of transition into the intermediate structure and from the intermediate structure to the denatured form were also contrasting. Interestingly, however, the transition temperatures of each transition remained relatively constant. This indicates that the increase in positive surface charge and subsequent alteration to the electrostatic stabilising forces present on the surface of C-PK had a favourable impact to the structure of PK and also altered the thermodynamic properties of the protein. Due to the simplicity of the model used to calculate the unfolding thermodynamics and the complexity of a three-state unfolding protein, it cannot be stated without

further investigation whether cationization has a positive or negative impact on the thermodynamics of PK, merely that they are altered in a variety of ways.

Previous investigations into the solvent-free liquid protein synthesis have shown similar results as displayed here, the effect of cationization being protein dependent^{6,7,9,10}. However, no matter the effect of cationization on the structure of the protein in question, the electrostatic conjugation of anionic surfactant molecules to the surface has consistently returned some native-like properties to the cationized protein and have subsequently have all been successfully transformed into solvent-free liquid proteins.

The following section examines how the addition of surfactant molecules to the surface of C-BSA to form aqueous [C-BSA][S] and to the surface of C-PK to form aqueous [C-PK][S] alters the physical characteristics of the proteins in relation to the native and cationized forms.

3.2.2 Characterisation of surfactant nanoconjugates

The second stage of the solvent-free liquid protein synthesis is the formation of surfactant-protein conjugates through the electrostatic conjugation of anionic surfactant molecules (Chapter 2.3.2) to the now highly cationized surface of the proteins in question. This is followed by exhaustive dialysis to remove any unbound surfactant. The surfactants used were termed S_1 and S_7 as detailed in Chapter 2, both displaying an anionic head group, SO_3^- and COO^- respectively. Negatively charged surfactants, such as sodium dodecyl sulphate (SDS), are often used in conjunction with proteins to allow for the separation of proteins and the determination of the protein's molecular weight by the complete denaturation of the protein through electrostatic and hydrophobic interactions. Since it would be detrimental to the overall aim for the proteins to unfold once conjugated it is important to investigate how these surfactants interact with C-BSA and C-PK. It is known from previous work^{6,7,10} that conjugation can initiate some reclamation of a native-like structure, hypothesised to do so through the shielding of the numerous intramolecular positive surface charges from one another. Even partial charge screening *via* interactions with the anionic head group would reduce the negative effect of like charge repulsion and could bring back some secondary structure characteristics that are removed during cationization. The presence of negative charges in areas that were previously negatively charged could create pseudo ion pairs, mimicking the ion pair that existed there before cationization, triggering further structure reclamation. This hypothesis does however also infer that areas that were originally positively charged will then have the presence of a negative charge, again creating a pseudo ion pair, potentially having a negative impact on the structure. This overall regaining of structure could be vital for the retention of reclamation of protease activity, as any alteration to the structure could lead to retardation or removal of protease activity towards the substrate.

To analyse how this surface modification altered the characteristics of the proteins in question, PK and BSA, measurements were taken to calculate the surface charge, the hydrodynamic radius, secondary structure, and thermal stability, as were conducted with the cationized species.

3.2.2.1 Charge density measurements using zeta potential

ZP measurements for aqueous [C-PK][S_1] and aqueous [C-BSA][S_1] gave values of -9 and -11 mV respectively, being more negative than the values for C-PK and C-BSA, +12 and +10 mV respectively. This shift in ZP is indicative of a surface charge alteration, from predominantly positive to predominantly negative. This suggests that the conjugation of surfactants to the surface has been successful, with the surfactant-protein conjugate displaying a charge density that is more representative of the negatively charged surfactant than that of the cationized protein beneath.

3.2.2.2 Particle size and distribution through dynamic light scattering (DLS)

DLS measurements of aqueous [C-BSA][S₁] and aqueous [C-PK][S₁] gave values for the hydrodynamic radii of 4.2 and 3.8 nm respectively (Figure 3.14, and Figure 3.15). Exhaustive dialysis was carried out to remove unbound surfactant.

Cationization of BSA, as discussed previously, caused the BSA to partially unfold and aggregate into positively charged clusters of C-BSA. Remarkably, the addition of the surfactant to the cationized aggregates clearly solubilised these aggregates, as evidenced by the reduction in the hydrodynamic radius (from 24.4 to 4.2 nm). It should be noted that the samples are filtered using a 0.45 µm filter, meaning it is unlikely that any surfactant conjugated aggregates that may have been present at a size smaller than this were removed through filtration.

The addition of surfactant to these aggregates could have a number of different effects. The surfactant could solely interact with the externally presented charges, effectively stabilising the aggregate and forming a small nanoconjugate protein cluster, not too dissimilar to the nanoclusters used in the production of enzymatically active self-standing protein-polymer surfactant films²². If this were the case, the radius of hydration of the aqueous [C-BSA][S₁] clusters would be similar to that of the C-BSA clusters, being 24 nm. As discussed previously, this is only double the radius of hydration of the protein haemoglobin, a protein that has displayed the capacity to be formed into a solvent-free liquid protein. It is also possible for the hydrophobic sections in the surfactant tail to interact with the hydrophobic interface, potentially stabilising partial aggregates. Another possibility once the surfactant begins conjugating with the C-BSA clusters is that the charge screening and formation of pseudo ion pairs begins to refold the protein, burying the hydrophobic core residues and removing the inter-protein β-sheets that lead to the aggregation of C-BSA. As the hydrophobic interactions are reduced the aggregates would disperse, with the surfactant subsequently interacting with the positively charged surface areas that were concealed by being internalised with the aggregate. This would lead to the formation of individual protein-surfactant nanoconjugates rather than clusters of nanoconjugates, having a smaller radius of hydration in comparison, making the formation of a solvent-free liquid BSA a more probable explanation.

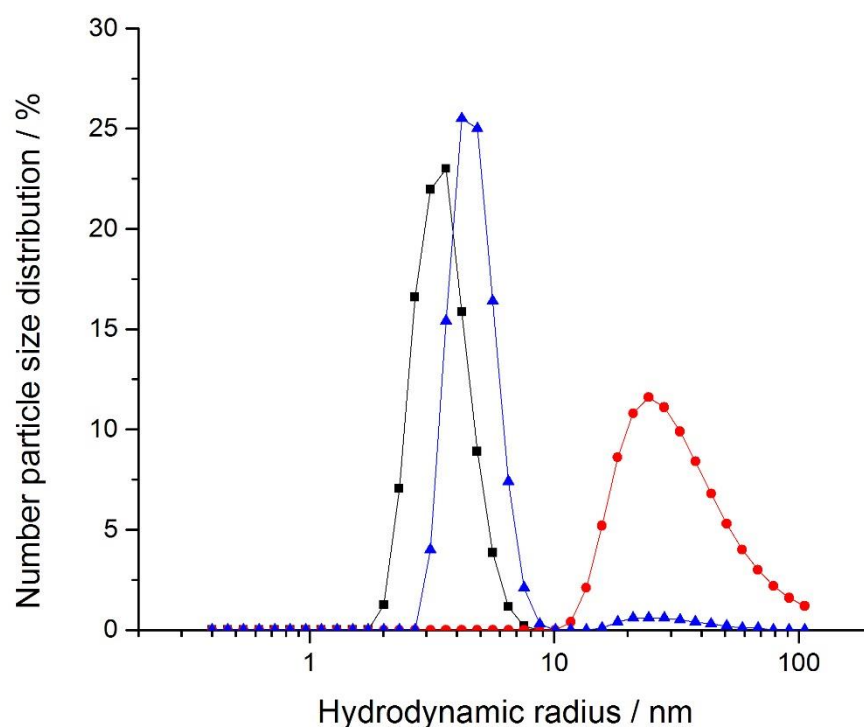


Figure 3.14: The number particle size distribution for aqueous BSA (black), aqueous cationized BSA (red), and aqueous S₁BSA (blue). [Conditions: unbuffered solutions at pH values of 6-6.5, with proteins concentrations of 0.5 mg.mL⁻¹, 25 °C]. Data displayed is an accumulation of at least 10 scans.

The conjugation of S₁ to C-BSA forming aqueous [C-BSA][S₁] has disrupted the C-BSA aggregates and has stabilized C-BSA in individual surfactant-protein units, as is clear by the large reduction in the 24 nm peak relating to the clusters and the formation of a 4.2 nm peak relating to individual aqueous [C-BSA][S₁] units. It is evident that some of these aggregates do still remain as the relating peak is non-zero but it is also evident that the majority has been disassembled. Relative to native-BSA, aqueous [C-BSA][S₁] has increased by 0.6 nm, a similar radius to that of S₁ micelles. Previous work indicated that the method of interaction between the micelles and the cationized protein was one of rapid envelopment²³. The negatively charged micelles are electrostatically attracted to the positively charged cationized proteins and, once in contact, the micelles deforms and the anionic head groups of the surfactants interact with positive charges on the surface of the protein. It is interesting to note that the average size of the aggregates has increased by ≈1 nm, displaying that the remaining aggregates have also been conjugated with surfactant molecules, yet this has not instigated the dispersion of these aggregates. How the surfactant conjugation altered the secondary structure relative to C-BSA and BSA is discussed in the following section.

Cationization of PK, as discussed previously, caused the secondary structure to increase relative to PK (0.8 nm increase) and to alter the thermal stability through the formation of a different less-stable unfolding intermediate species. This size increase was not coupled with aggregation and formation of

cationized-protein clusters as occurred with C-BSA. Similar studies with other protein systems suggest that the addition of surfactant molecules to C-PK forms individual surfactant-C-PK units^{23,24}. A plausible explanation is that the conjugation has initiated a small degree of structure reclamation where the modified PK has regaining native-like structure and size. Structure analysis is required to confirm this as discussed in the following section.

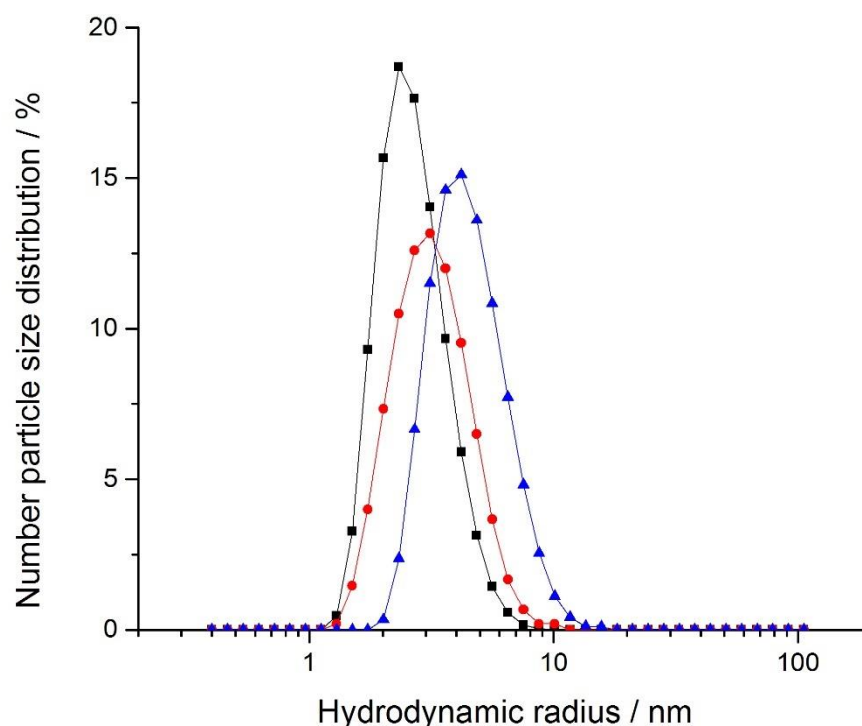


Figure 3.15: The number particle size distribution for aqueous PK (black), aqueous cationized PK (red), and aqueous S₁PK (blue). [Conditions: unbuffered solutions at pH values of 6-6.5, with proteins concentrations of 0.5 mg.mL⁻¹, 25 °C]. Data displayed is an accumulation of at least 10 scans.

The conjugation of S₁ to C-PK forming aqueous [C-PK][S₁] caused the average hydrodynamic radius to increase to 3.8 nm, a 0.7 nm increase relative to C-PK (Figure 3.15). The lack of peaks greater than 15 nm displays that the surfactant conjugation did not initiate unfolding and subsequent aggregation of aqueous [C-PK][S₁] units to form large aqueous [C-PK][S₁] clusters. This would indicate that C-PK successfully associated with the surfactant molecules to form individual aqueous [C-PK][S₁] units of a size that make the formation of a solvent-free liquid PK highly probable. The following section will discuss how the surfactant conjugation altered the secondary structure relative to C-PK and PK.

3.2.2.3 Secondary structure analysis through circular dichroism spectroscopy (CD)

3.2.2.3.1 Surfactant selection for CD spectroscopy

As previously detailed in Section 2.1.1, CD spectroscopy is a technique that can be utilised to study the secondary structure of a protein, with the spectra being deconvoluted in order to estimate the percentages of secondary structure elements. The technique is used here to study the effect of surfactant conjugation on the secondary structure of aqueous [C-BSA][S] and aqueous [C-PK][S] nanoconjugates relative to the native and cationized species of each protein. With this in mind, the secondary structure analysis for the surfactant conjugated BSA and PK species was performed using straight chained S_7 as the surfactant. This was used because the S_1 surfactant strongly absorbs at wavelengths below 300 nm due to the aromatic ring present in the polymer tail, and as secondary structure data from CD is obtained between 190-260nm the data would be unreliable and noisy. S_7 is similar to S_1 in many ways (structures displayed in Section 2.3.1.1): the head group has a negative charge, there is a repeating PEG section of the chain following the head group and a long hydrophobic chain following the PEG section. The differences between the two surfactants are: the negative head group of S_7 is formed from COO^- in contrast to the SO_3^- of S_1 , there are an additional 3 PEG units on S_7 , and there is no aromatic ring between the PEG units and the long carbon chain, with the carbon chain being 3 carbons longer as well. Therefore, to avoid the potential spectroscopic issues and with assurances of similar physical characteristics due to its precedence in previous work, the S_7 surfactant was used in place of S_1 .

3.2.2.3.2 Structure analysis of aqueous [C-BSA][S_7] nanoconjugates

The formation of aqueous [C-BSA][S_7] decreases the intensity of the negative 208 nm peak relative to C-BSA, and also red-shifts the parallel $\pi \rightarrow \pi^*$ transition (205-208 nm) and the $\pi \rightarrow \pi^*$ transition (190-193 nm), returning the characteristic peaks to wavelengths similar to that of native-BSA (Figure 3.16). This demonstrates that the cause for the blue shifts in CD spectra of the protein, whether that was loss of structure or another factor brought about due to cationization, was corrected for by the conjugation of surfactant to the positive surface sites. The reduction in intensity of the 208 nm peak relative to C-BSA shifts the $\theta_{222}/\theta_{205}$ ratio to ≈ 1 , similar to that of the native-BSA, suggesting that the 3_{10} helices were stabilized by the interactions with the surfactant to re-form α -helices¹². However, as the 194 nm peak remains similar to the other two characteristic peaks, it is unlikely that there is a large reclamation of α -helical structure within the protein. The CD spectrum of aqueous [C-BSA][S_7] displays potential structure reclamation, in the form of red-shifting characteristic peaks, although the intensity of aqueous [C-BSA][S_7] is still lower than that of BSA, being more comparable to that of C-BSA.

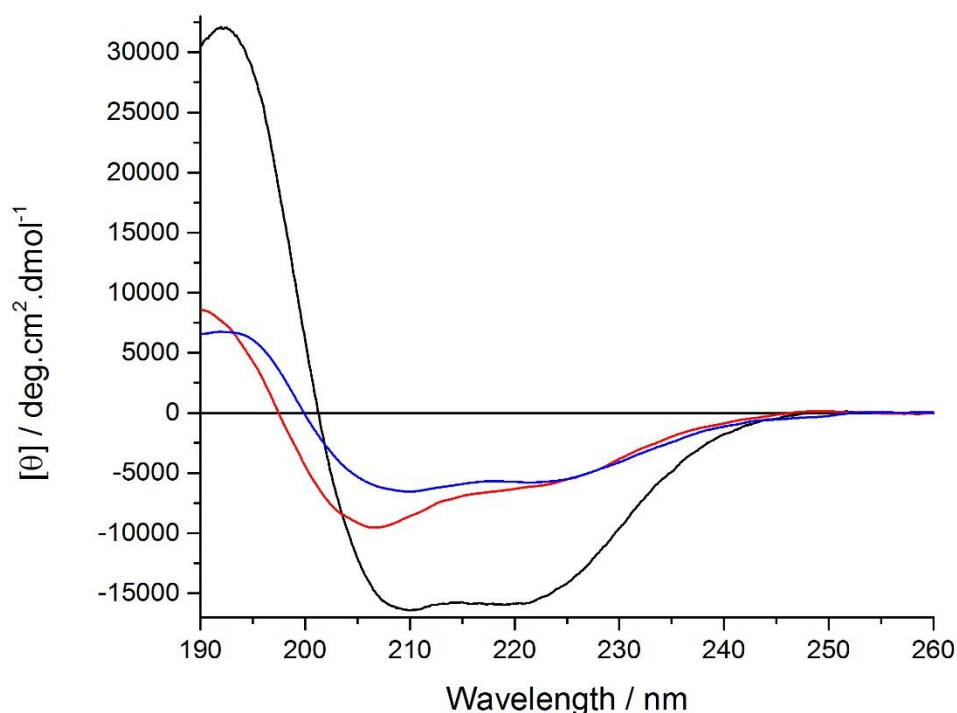


Figure 3.16: Mean residue ellipticity against wavelength showing far-UV CD spectra of aqueous BSA (black), C-BSA (red), and aqueous [C-BSA][S₇] (blue) at 25 °C between 190-260nm.

It is clear from the deconvoluted data that the surfactant conjugation of C-BSA to form aqueous [C-BSA][S₇] marginally altered the secondary structure, with a limited decrease to the β -sheet content whilst increasing the unordered content, an overall change of 4% for each (Table 3.8). This implies that, contrary to previous solvent-free liquid protein work¹⁰, the addition of surfactant molecules initiates destabilisation of β -sheet content with no increase in α -helical content, although bringing the β -sheet content back toward the native structure. This loss of β -sheet structure provides further explanation towards the reduction in aggregation, dispersing the aggregates due to lack of inter-protein β -sheets, forming molecularly dispersed BSA-surfactant molecules. Observations such as these indicate that the screening of 100% of the surface charges through addition of anionic surfactant does not always return the structure to native-like conformation, showing that the presence of some surface repulsion aids in the stabilisation of the conformation. Therefore, the removal of all repulsive events has an overall minor negative effect, with the possibility of some structure reclamation being overshadowed by a larger loss in structure.

Table 3.8: Table of secondary structure content calculated from CD spectra displayed in Figure 3.16.

Sample	α -Helix / %	β -Sheet / %	Turns / %	Other/ %
Aqueous [C-BSA][S ₇]	18	22	14	46
Aqueous C-BSA	18	26	14	42
Aqueous native-BSA	37	14	12	37

3.2.2.3.3 Structure analysis of aqueous [C-PK][S₇] nanoconjugates

The formation of aqueous [C-PK][S₇] reduced the intensity of all three characteristic peaks, broadening the 222 and 208 nm peaks, and dropped the $\theta_{222}/\theta_{208}$ ratio to 1.37 (Figure 3.17). As well as a broadening of the 222 and 208 nm peaks, there is a red-shift of the $n \rightarrow \pi^*$ transition peak (222-223 nm) and of the positive $\pi \rightarrow \pi^*$ transition peak (>190 - 190 nm), along with a blue shift of the parallel $\pi \rightarrow \pi^*$ transition peak (208-207 nm) (Figure 3.17). Similar to the surfactant conjugation of BSA, the change in intensity is not even across the characteristic peaks, with a significant decrease in the ellipticity of the positive 190 nm peak, along with a larger decrease in the negative 222 nm peak relative to the negative 210 nm peak, decreasing the peak ratio from 1.5 to 1.37. A decrease this large in the positive 190 nm peak alongside a reduction in the 210 nm peak indicates a reduction in perpendicular and parallel $\pi \rightarrow \pi^*$ transitions which can signify a loss of α -helical structure. The decrease in intensity of the positive $\pi \rightarrow \pi^*$ transition and the $n \rightarrow \pi^*$ could also indicate a loss of β -Sheet structure. With a decrease in all characteristic peaks, it is also likely that there is an increase in random coil percentage, a decrease in signal also signifying an increase in the opposite transition.

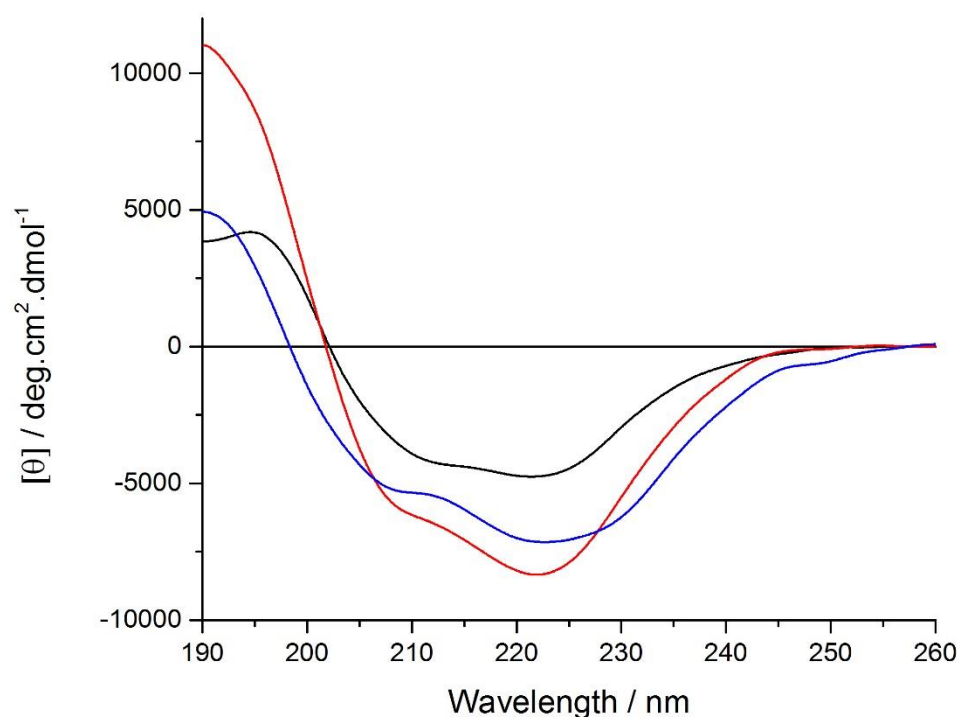


Figure 3.17: Mean residue ellipticity against wavelength showing far-UV CD spectra of aqueous PK (black), C-PK (red), and aqueous [C-PK][S₇] (blue) at 25 °C between 190-260nm.

It is clear from the deconvoluted data that the surfactant conjugation of C-PK to form aqueous [C-PK][S₇] negatively altered the secondary structure, with a decrease to both the α -helical and β -sheet content, 8% and 7% reduction respectively, whilst increasing the turn and random coil content (Table 3.9). The return of the α -helical content to native-like levels indicates that the screening of surface

charges through the addition of surfactants may potentially have some positive impact on the active conformation of the protein. However, a return to the same calculated α -helical value does not confirm a return to native α -helical structure, merely that the percentage of the structure present has the same α -helical content. Therefore, as the active site of PK is known to be proximal to α -helices, a shift to a different α -helical structure could again alter the activity. In addition, a reduction of the core β -Sheet content may imply a more drastic conformational change, an extended β -Sheet being a core PK characteristic, again potentially altering the activity of aqueous [C-PK][S₇] relative to C-PK and native-PK.

Table 3.9: Table of secondary structure content calculated from CD spectra displayed in Figure 3.17.

Sample	α -Helix / %	β -Sheet / %	Turns / %	Other/ %
Aqueous [C-PK][S ₇]	10	26	17	46
Aqueous C-PK	18	33	13	37
Aqueous native-PK	10	34	13	43

3.2.2.4 Thermal denaturation

3.2.2.4.1 Thermal denaturation of aqueous [C-BSA][S₇] nanoconjugates

The [C-BSA][S₇] nanoconjugate exhibited a more sigmoidal-like 222 nm denaturation curve, relative to C-BSA, with clear pre-transition and post-transition plateaus (Figure 3.18B). This suggests that there is a degree of reclamation of the cooperative unfolding mechanism displayed in native-BSA, however, the noise in the spectrum along with the shallower transition may indicate retention of some molten globule-like characteristics. The reclamation of a cooperative unfolding mechanism once the surfactant is conjugated could demonstrate a screening of additional surface charges reducing the repulsive forces, limiting the molten-globular effect and forming a relatively more densely packed protein. It could also be due to the reduction in aggregation once the surfactant is bound, implying that it is individual [C-BSA][S₇] nanoconjugates being investigated at the start of the thermal test rather than a cluster of nanoconjugates. It is likely that the unfolding dynamics of a protein aggregate vastly differ from that of an individual protein due to the large increase in intramolecular interactions, potentially forming a vast array of secondary structures throughout the unfolding mechanism that can stabilize the protein in a different way. The deconvoluted data supports this assertion further by displaying a drastic shift in secondary structure motifs, starting at 30 °C and ending at 80 °C, where the α -helical content reaches 0% (Figure 3.18D) and the β -sheet content plateaus at a maximum value. As the α -helical content decreases there is a commensurate increase in observed β -sheet content, displaying that the α -helices unravel and form β -sheet structures, both internally and potentially externally through inter-protein β -sheets. Interestingly the turn and unordered content remain

relatively constant throughout, suggesting that the turns are relatively more resistant to thermal change.

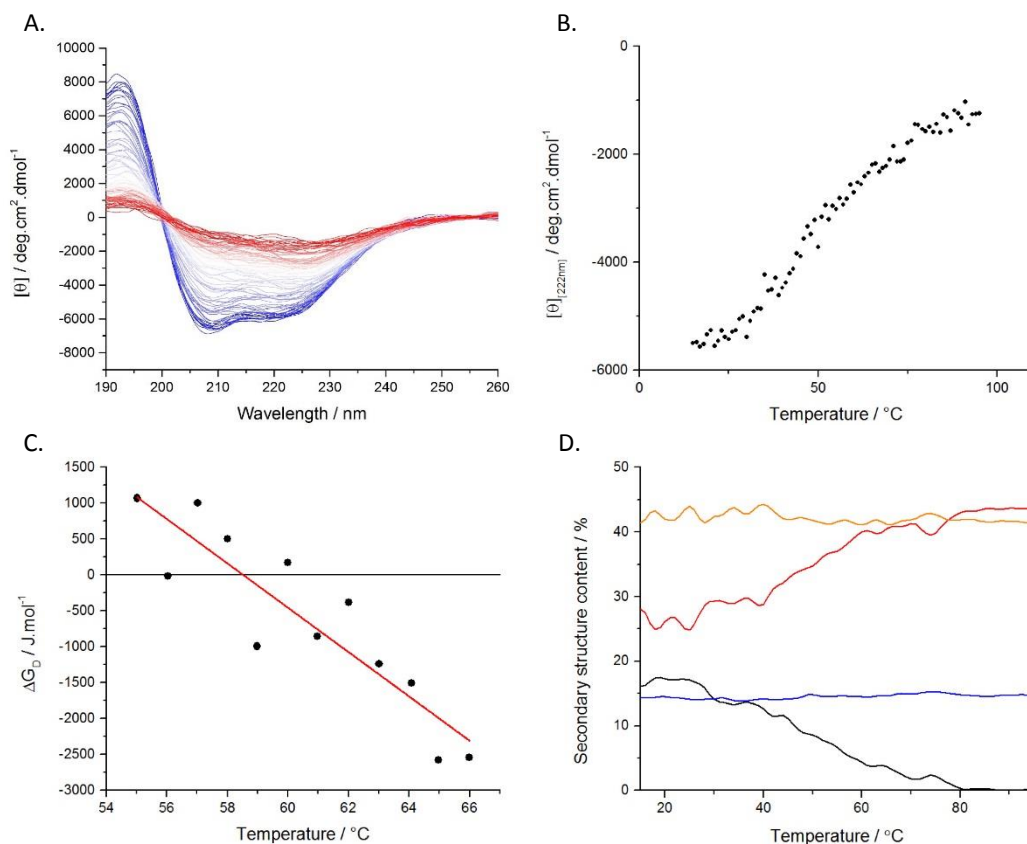


Figure 3.18: **A.** Mean residue ellipticity against wavelength showing far-UV CD spectra for aqueous [C-BSA][S₇], with the gradient from blue to red indicating an increase in temperature from 20 – 95 °C, **B.** Mean residue ellipticity at 222 nm against temperature (20 – 95 °C) for aqueous [C-BSA][S₇], **C.** plot of free energy of denaturation (ΔG_D) against temperature displaying the transition region for aqueous [C-BSA][S₇] (the linear fit (red) was used to calculate the thermodynamic parameters), and **D.** secondary structure percentage change against temperature (20 – 95 °C) for aqueous [C-BSA][S₇] (secondary structure percentage calculation detailed in Chapter 2), with the percentage of α -helices (black), β -sheets (red), random coil (orange), and turns (blue).

The model used to calculate T_m can be appropriately implemented here due to the cooperative unfolding of the [C-BSA][S₇] nanoconjugate. The melting transition was calculated to be 59 °C; 30 °C lower than the estimated C-BSA T_m , but interestingly only 13 °C lower than native-BSA. As the entropy and enthalpy values for the transition are approximately 1/3 of the value for the native-BSA transition (Table 3.10) this shows that the energy released when unfolding is significantly lower than native-BSA. This is down to the entropy of the system already being lower than that of native BSA due to the surface modification; [C-BSA][S₇] nanoconjugate is in a lower entropic state than native BSA and therefore is more stable to the temperature change. There is a large error associated with each calculated value (T_m , ΔH_m , and ΔS_m), which supports the theory that there is retention of some molten-globular characteristics. This again indicates that whilst there is the reclamation of native-like unfolding in the form of a two-state unfolding mechanism, the structure is already in a much more unfolded state; it takes a lower energy to form the denatured conformation.

Table 3.10: Table displaying the transition melting temperature, enthalpy change for melting transition, and entropy change for melting transition (T_m , ΔH_m , and ΔS_m respectively) for aqueous BSA, C-BSA, and [C-BSA][S₇]. *calculated through chemical denaturation.

Sample	$T_m / ^\circ\text{C}$	$\Delta H_m / \text{kJ mol}^{-1}$	$\Delta S_m / \text{J K}^{-1} \text{mol}^{-1}$
Aqueous [C-BSA][S ₇]	59 ± 12	102 ± 16	309 ± 46
Aqueous C-BSA	89 ± 1 *	N/A	N/A
Aqueous native-BSA	72 ± 3	320 ± 14	930 ± 9

3.2.2.4.2 Thermal denaturation of aqueous [C-PK][S₇] nanoconjugates

It is clear from the C-PK thermal denaturation CD curve at 222 nm (Figure 3.19) that the surfactant conjugation altered the thermal unfolding dynamics relative to C-PK, although, again the double sigmoidal characteristic was retained indicating that an intermediate species is still formed. A key difference when comparing aqueous [C-PK][S₇] with C-PK and PK is that the intermediate species that is formed doesn't remain above 50 °C, showing that, as with C-PK, the species formed is different to and less thermally stable than the species formed with PK. Interestingly, the shape of the denaturation curve of aqueous [C-PK][S₇] resembles that of PK more than that of C-PK. This indicates that there is a degree of reclamation of the native unfolding pathway, with the intermediate stabilising at a lower temperature. Despite this difference, the temperatures of transition, T_{m1} and T_{m2} , are similar in value those of C-PK, being 39 and 59 °C respectively, with T_{m2} also being similar to that of PK. As with PK, both entropies and enthalpies of the transitions are similar, indicating that the intermediate species is approximately as stable as the folded species (Table 3.11). However, both entropies and enthalpies of the transitions are lower than that of PK, meaning that there is less energy released upon the denaturation of the species due to a smaller rearrangement of conformation. The lower entropy and enthalpy for the T_{m2} of aqueous [C-PK][S₇] relative to C-PK indicates that the intermediate species of C-PK is more thermally stable than that of aqueous [C-PK][S₇].

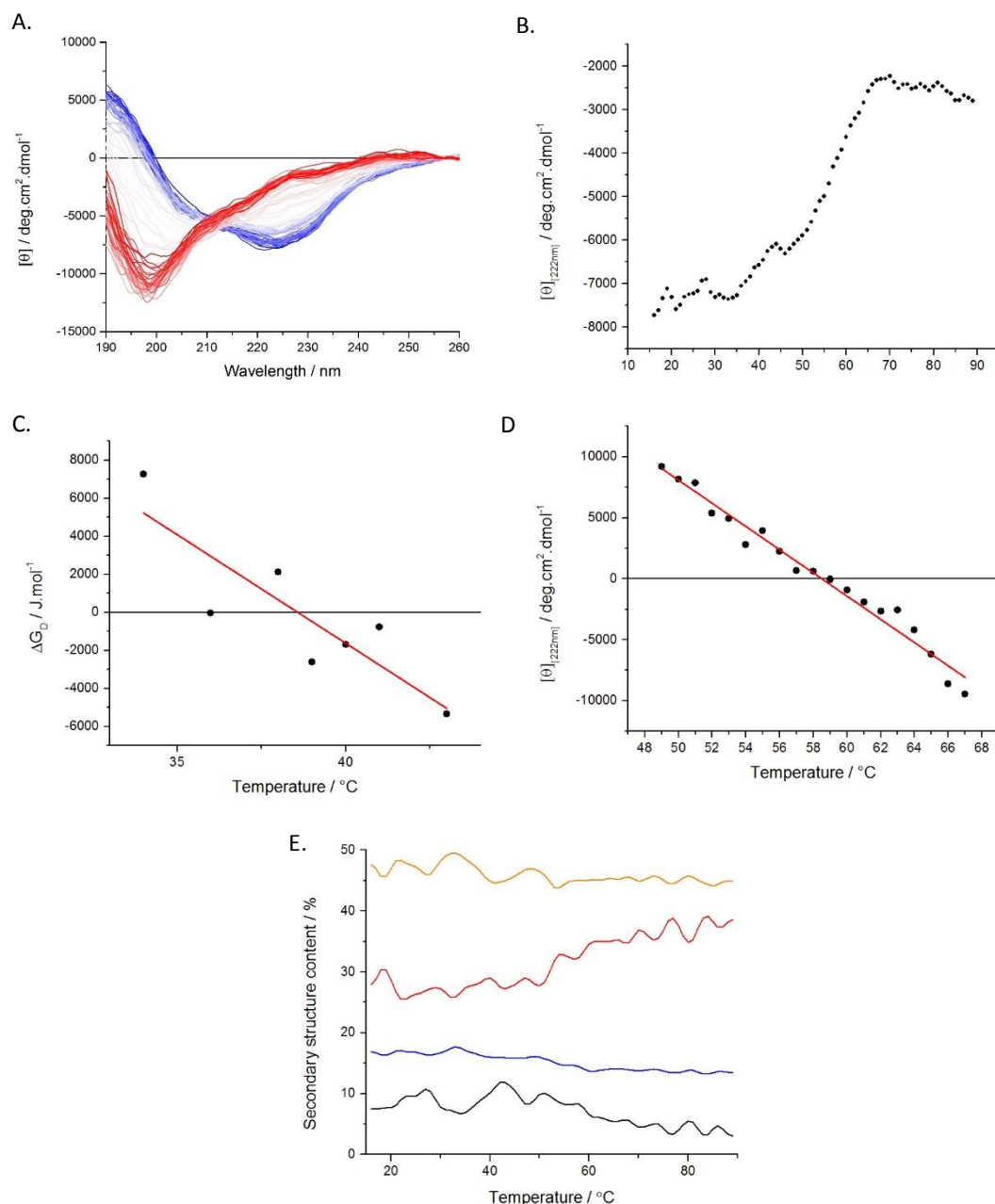


Figure 3.19: **A.** Mean residue ellipticity against wavelength showing far-UV CD spectra for aqueous [C-PK][S₇], with the gradient from blue to red indicating an increase in temperature from 20–95 °C, **B.** Mean residue ellipticity at 222 nm against temperature (20–95 °C) for aqueous [C-PK][S₇], **C.** plot of free energy of denaturation (ΔG_D) against temperature displaying the transition region for aqueous [C-PK_a][S₇] (the linear fit (red) was used to calculate the thermodynamic parameters), **D.** plot of free energy of denaturation (ΔG_D) against temperature displaying the transition region for aqueous [C-PK_b][S₇] (the linear fit (red) was used to calculate the thermodynamic parameters), and **E.** secondary structure percentage change against temperature (20–95 °C) for aqueous [C-PK][S₇] (secondary structure percentage calculation detailed in Chapter 2), with the percentage of α -helices (black), β -sheets (red), random coil (orange), and turns (blue).

Table 3.11: Table displaying the transition melting temperature, enthalpy change for melting transition, and entropy change for melting transition (T_m , ΔH_m , and ΔS_m respectively) for aqueous PK, C-PK, and [C-PK][S₇].

Sample	$T_m / ^\circ\text{C}$	$\Delta H_m / \text{kJ mol}^{-1}$	$\Delta S_m / \text{J K}^{-1} \text{mol}^{-1}$
Aqueous [C-PK][S ₇] b	59 ± 3	320 ± 12	950 ± 38
Aqueous [C-PK][S ₇] a	39 ± 14	350 ± 89	1100 ± 280
Aqueous C-PK b	61 ± 10	590 ± 73	1800 ± 220
Aqueous C-PK a	39 ± 4	220 ± 15	700 ± 49
Aqueous native-PK b	61 ± 6	410 ± 29	1200 ± 87
Aqueous native-PK a	33 ± 16	450 ± 160	1500 ± 500

3.2.2.5 Conclusions for surfactant conjugation

The second stage of the solvent-free liquid protein synthesis is surfactant conjugation; the electrostatic addition of anionic polymer surfactant molecules to the surface of the highly cationized protein. Surfactant addition is required to form the necessary steric coronal layer around the protein, increasing the interaction distance between two interacting proteins and allowing a liquid-protein phase to form. In the synthesis of previously formed solvent-free liquid proteins the surfactant addition stage had a number of effects: the displayed surface charge of the cationized protein was altered; a marginal increase in the hydrodynamic radii; often the secondary structure of the cationized protein returned to more native-like conformation; and the thermodynamics of unfolding were changed.

Surfactant addition to both C-BSA and C-PK, forming aqueous [C-BSA][S₁] and aqueous [C-PK][S₁], had a similar effect regarding surface charge as seen in the synthesis of previous solvent-free liquid proteins. Through electrostatic attraction, negatively charged surfactant and the cationized proteins were attracted and interacted. Once the polymer had interacted with the positively charged surface sites, the polymer nanoconjugates displayed a negative surface charge, thought to be due to the anionic head groups, the hydrophilic PEG units being displayed externally and the hydrophobic units wrapping closely to the surface of the protein.

C-BSA, as detailed previously, formed aggregates due to the exposure of hydrophobic core residues, increasing the average hydrodynamic radii. After surfactant addition the average hydrodynamic radii decreased from 24 nm to 4.2 nm, implying a disbandment of the aggregates, and therefore a reclamation of individual protein constructs, due to the covering of the hydrophobic residues by the hydrophobic section of the surfactant tail. This removed the need for the aggregate stabilising inter-protein β -sheets, shown experimentally by a reduction in β -sheet content and therefore some native-structure reclamation. The radii of both aqueous [C-BSA][S₁] and aqueous [C-PK][S₁] was ≈ 1 nm larger than the respective native, with the increase in size being mainly attributed to the additional surfactant length with the potential of a slight increase in protein size due to altering surface interactions.

The 25 °C CD spectrum of aqueous [C-BSA][S₇] displayed only a minor alteration to secondary structure relative to C-BSA, with a similarly minor change to estimated secondary structure percentages. The $\theta_{222}/\theta_{205}$ peak ratio returned to ≈ 1 from 0.69, indicating an alteration to the secondary structure. There

was also a shift in characteristic peak maxima/minima, returning them to more native-like wavelengths. This shows that there is a small degree of structure reclamation, reverting the secondary structure to more native-like. Surfactant addition partially reclaimed the sigmoidal nature of the BSA thermal denaturation curve, again indicating a small degree of native-like structure reclamation. The T_m of aqueous [C-BSA][S₇] was 13 °C lower and the entropy and enthalpy of transition were both 2/3rds lower than that of native-BSA, indicating that the [C-BSA][S₇] nanoconjugate is in a lower entropic state than that of native BSA. Therefore, it has been demonstrated that surfactant conjugation of C-BSA had a smaller effect on the surface of the secondary structure relative to that observed in previously developed solvent-free liquid proteins; there being no appreciable native-like structural reclamation. Although, it is important to emphasize that the surfactant addition did cause the majority of the aggregates to dissolve.

Analysis of the 25 °C CD spectrum of aqueous [C-PK][S₇] revealed a large alteration in secondary structure and the estimated secondary structure percentages, demonstrating a return to a more native-like structure relative to C-PK and native-PK. In contrast with aqueous [C-BSA][S₇], the $\theta_{222}/\theta_{205}$ peak ratios were lowered from 1.52 to 1.37, were large decreases in the 190 nm and 222 nm characteristics peaks implying a decrease in α -helical content. This reversed the some of the effects of PK cationization; decreasing the secondary structure and bringing the estimated percentage of α -helices back to 10% providing evidence for structure reclamation. This is important to note due to the location of the protease active site being close in proximity to several α -helices, where any alteration could impact upon the activity of the protease.

Thermal degradation of aqueous [C-PK][S₇] again displayed retention of the double sigmoidal thermal degradation pathway, with the curve having more native-like characteristics relative to C-PK; the intermediate species being stabilized at a temperature closer to that of native PK than of C-PK. As with C-PK, the intermediate species differs to that of native-PK and is also not stable above 50 °C, having a small stable temperature range compared to the ≈ 15 °C range for native-PK and C-PK. The calculated T_m of the transition to the intermediate state (T_{ma}) was identical to C-PK, 39 °C, being 6 °C higher than native-PK. The entropy and enthalpy of this transition were lower than that of the native-PK, indicating that the transition to the intermediate state required a lower entropy change and therefore lower energy. This, along with the differing stabilized ellipticity, provides evidence for the intermediate species again being different to that of native-PK. T_{mb} was marginally lower than for native-PK and C-PK, 59 °C rather than 61 °C, indicating that the intermediate species is not as stable relative to the other two intermediate species. This is also reflected in the entropy and enthalpy changes of this transition to the denatured state, being appreciably lower than the other two species.

In conclusion, the electrostatic addition of anionic surfactant molecules to the surface of a highly positively-charged protein has varying effects based on the specific protein being surface modified. A common effect found in previous solvent-free liquid protein syntheses is that the surface charge increases and there is a marginal increase in the hydrodynamic radii relative to the native species. Each

protein modified in this way has also displayed an alteration to the secondary structure, bringing the structure of the cationized protein back to a more native-like conformation. It has been hypothesised that the mechanism responsible for this is the screening of the numerous same charge surface interactions present due to cationization. The anionic head group of the surfactant interacts with these point charges and removes the cumulative negative effect, restoring some degree of structure. However, the screening of all charges (with each individual surfactant interacting with one complimentary cationized surface site) may also shed light on why the native structure is not fully restored; with a degree of surface repulsion being intrinsic for the native structure to exist. Alongside repulsive surface charges, the native structure would also have had a number of surface salt bridges that are disrupted through cationization, with these not being restored by the addition of surfactant. As these surface repulsions and additional surface interactions have not been replaced, the structure was not fully reclaimed for either protein in question. Molecularly dispersed protein-polymer surfactant nanoconjugates with significant levels of native secondary structure have been produced, indicating that the final stage of the synthesis can be carried out; the formation of solvent-free liquid proteins.

The following section examines how removing water and thermally annealing aqueous [C-BSA][S] to form solvent-free liquid BSA and of aqueous [C-PK][S] to form solvent-free liquid PK, alters the physical characteristics of the proteins in relation to the native, cationized and surfactant nano-conjugated forms of the proteins.

3.2.3 Characterisation of solvent-free liquid proteins

The final stages of the solvent-free liquid protein synthesis is lyophilization and thermal annealing, two stages that would form a dry powder or denature a non-modified protein. Due to the additional interactions from the surfactant molecules increasing the interaction distance between protein nanoconjugates, the removal of water and heating allows for the formation of a solvent-free liquid protein phase that has previously displayed a high retention of native-like protein secondary structure. This phase formation and retention of the structure is possible in the absence of water due to the essential water-protein hydrogen bonding interactions being replaced and replicated by the interaction of the surfactant with the protein. The investigation into the formation of a solvent-free liquid BSA and a solvent-free liquid PK is discussed in the following section.

Solvent-free liquid BSA and solvent-free liquid proteinase K were created for the first time. Once frozen in liquid nitrogen and subsequently lyophilised, the resultant solvent-free protein sample is a low density solid protein (Figure 3.20A) that once thermally annealed at 50 °C melts into a solvent-free liquid protein (Figure 3.20B). Only proteinase K is displayed as both the solvent-free low-density solids and the solvent-free liquids were indistinguishable from one another. The investigation into and characterisation of solvent-free liquid BSA and solvent-free liquid PK is discussed in the following sections.

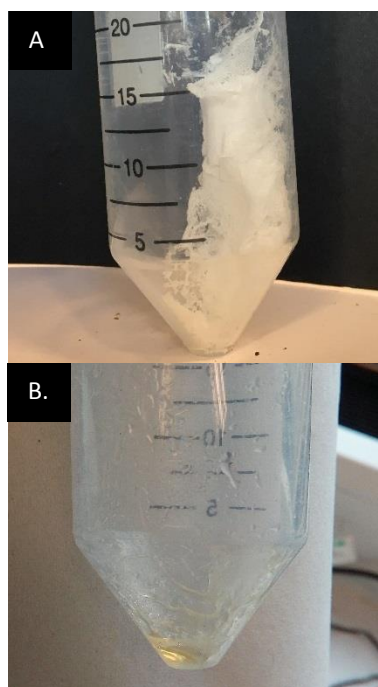


Figure 3.20: Images displaying **A.** solvent-free proteinase K as low-density solid prior to annealing and **B.** solvent-free proteinase K in a melted liquid form post-annealing at 50 °C.

3.2.3.1 Differential scanning calorimetry

Differential Scanning Calorimetry (DSC) is a thermo-analytical technique used to study the thermal properties of a material and has been described previously (Chapter 2). It has been demonstrated that solvent-free liquid proteins have a liquid phase that occurs at a protein-dependent temperature⁹. DSC was used to establish the nature of the melting transitions in the material created through the solvent-free liquid protein synthesis and to determine at what temperature these melting transitions occur. The DSC traces of each solvent-free liquid protein created were compared to the DSC traces of the surfactant used in order to establish whether the transition is unique to the liquid protein constructs or whether the transition was purely due to the surfactant.

3.2.3.1.1 DSC of solvent-free liquid [C-BSA][S₁]

Lyophilised proteins are known not to have a liquid phase (Figure 3.21), with the DSC trace of lyophilized BSA having no significant increases or decreases in heat flow over the temperature range of -70 – 200 °C. In contrast to lyophilized BSA, the DSC traces of the solvent-free liquids studied show both a reversible melting transition and crystallization transition peaks (Figure 3.22), indicating the ability of solvent-free liquid proteins to exhibit a reversible transition from the solid phase to the liquid phase.

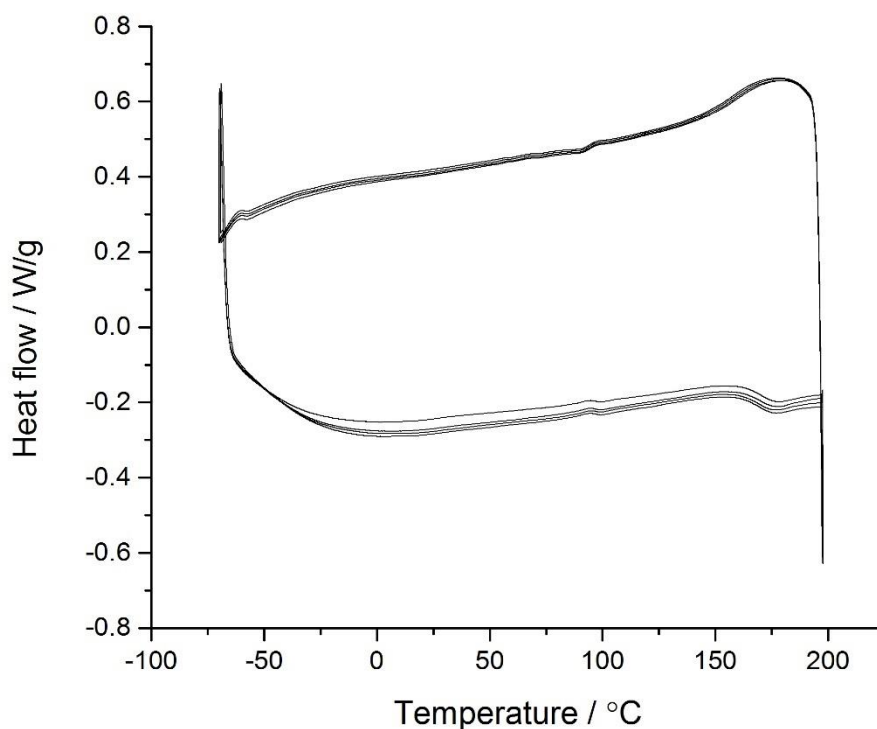


Figure 3.21: Cyclic DSC traces showing normalized heat flow against temperature for lyophilized BSA. [Experiments were cycled a minimum of 6 times and the scan rate was 10 °C.min⁻¹].

DSC traces of solvent-free liquid [C-BSA][S₁] displayed a relatively broad reversible exothermic crystallization peak at 2 °C and a sharp reversible endothermic melting transition peak at 26 °C in the heating cycle (Figure 3.22). These peaks were accompanied by observable glass transitions occurring at -51 °C in the heating cycle and at -50 °C in the cooling cycle. These thermal transitions are in contrast to the ones displayed on the DSC traces for S₁, both in value and shape. S₁ displayed a sharp reversible exothermic crystallization transition peak at -16 °C and sharp reversible endothermic melting transition peak at 32 °C in the heating cycle. These peaks were accompanied by an observable glass transition occurring at -47 °C in the heating cycle, a small broad exothermic transition peak at -20 °C and a second glass transition at -27 °C in the cooling cycle. For solvent-free liquid [C-BSA][S₁], the crystallization temperature was 18 °C higher and the melting transition temperature was 6 °C lower than that of S₁, along with a slight shift in glass transition temperatures. It is also observable that the transitions observed in S₁ have a higher heat flow change.

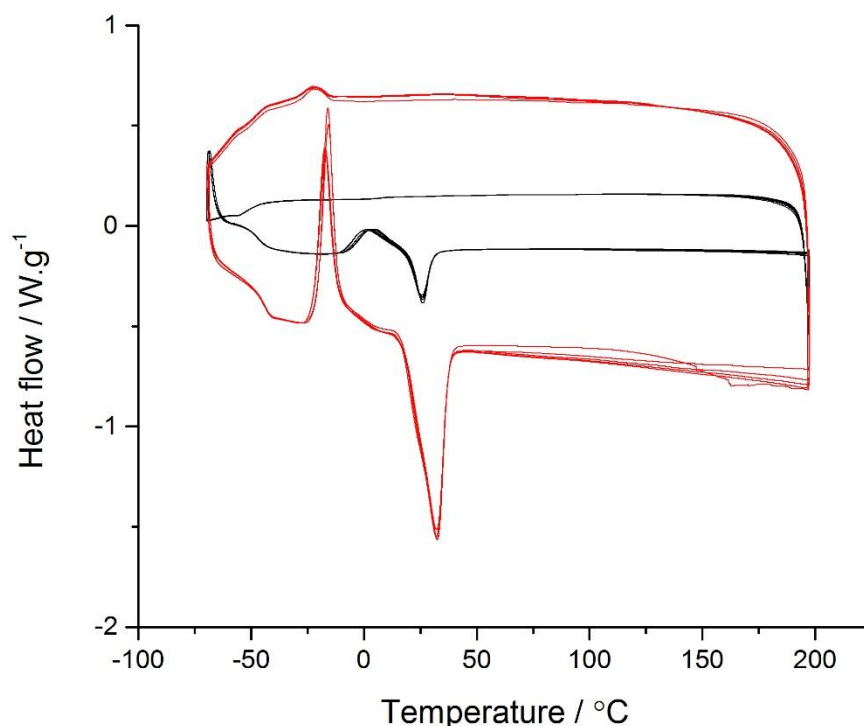


Figure 3.22: Cyclic DSC traces showing normalized heat flow against temperature for solvent-free liquid [C-BSA][S₁] (black) and S₁ (red). [Experiments were cycled a minimum of 6 times and the scan rate was 10 °C.min⁻¹].

From the DSC traces it is clear that the crystallization transition occurs within the heating cycle of the solvent-free [C-BSA][S₁] after it has passed through the glass transition. The thermal behaviour on the cooling cycle of S₁ differed to that of solvent-free [C-BSA][S₁], in that there is a small endothermic peak positioned at -21 °C on the cooling cycle, with the full crystallization transition occurring at a temperature of -16 °C on the heating cycle. This indicates that S₁ has partial crystallization on the cooling cycle and requires a lower crystallization activation energy than solvent-free [C-BSA][S₁].

The lack of a full crystallization transition on the cooling cycle in both S_1 and solvent-free [C-BSA][S_1] could be due to the scan rate being faster than the kinetics of the system, *i.e.* the crystallization time of domains within the materials are slower than that of the scan rate. The domain reorientation through the two glass transitions along with the temperature increase allows for the crystallization transition to occur within the heating cycle.

The difference in thermal properties, *i.e.* the shift in melting and crystallization transition temperatures, between S_1 and solvent-free liquid [C-BSA][S_1] demonstrate that solvent-free liquid [C-BSA][S_1] is a distinct material from S_1 formed of discrete particles, and not just BSA that has been dissolved within S_1 surfactant.

3.2.3.1.2 DSC of solvent-free liquid [C-BSA][S_7]

DSC traces of solvent-free liquid [C-BSA][S_7] displayed a sharp reversible exothermic crystallization peak at -5 °C in the cooling cycle and sharp reversible endothermic melting transition peak at 44 °C in the heating cycle; no glass transitions were observed (Figure 3.23). These thermal transitions are in contrast to the ones displayed on the DSC traces for S_7 , with a marginally different trace profile and differing transition temperatures. S_7 displayed a sharp mostly reversible exothermic crystallization transition peak at 1 °C in the cooling cycle and sharp reversible endothermic melting transition peak at 40 °C in the heating cycle. These peaks were accompanied by three observable glassy transitions occurring at -36 °C, -6 °C and 22 °C in the heating cycle and a fourth glassy transition at -60 °C in the cooling cycle. For solvent-free liquid [C-BSA][S_7], the crystallization temperature was 6 °C lower and the melting transition temperature was 4 °C higher than that of S_7 , along with the loss of the glassy transitions when solvent-free liquid [C-BSA][S_7] was formed.

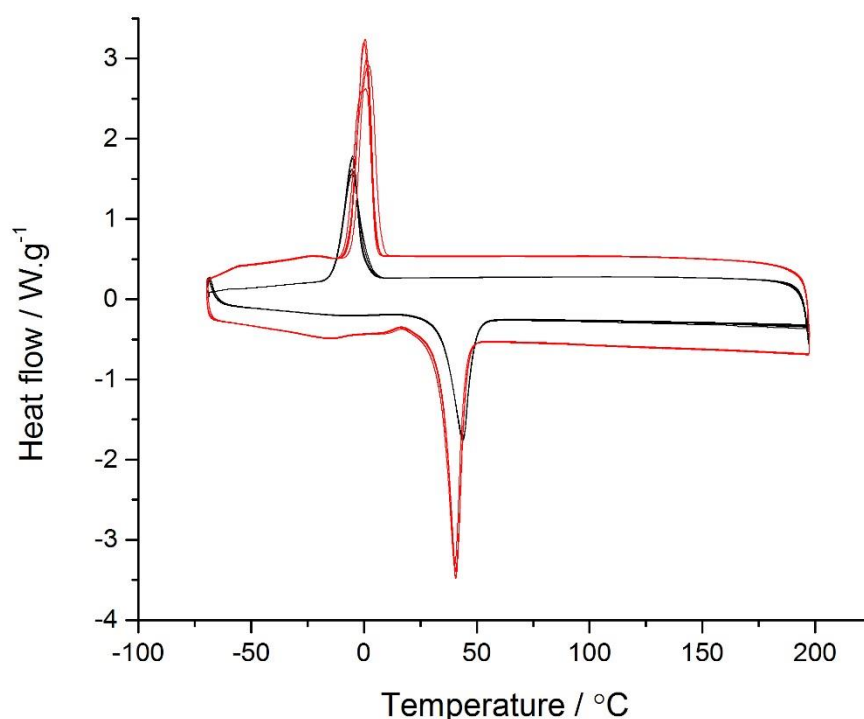


Figure 3.23: Cyclic DSC traces showing normalized heat flow against temperature for solvent-free liquid [C-BSA][S₇] (black) and S₇ (red). [Experiments were cycled a minimum of 6 times and the scan rate was 10 °C.min⁻¹].

From the DSC traces it is clear that the crystallization transition occurs within the cooling cycle of the solvent-free liquid [C-BSA][S₇]. This indicated that upon reaching the crystallization temperature of -5 °C the activation energy for crystallization is reached, with the crystallization transition being observed. The thermal behaviour of the cooling cycle of S₇ differed marginally to this, in that crystallization transition occurred 6 °C higher. This indicates that S₇ requires a lower crystallization activation energy than solvent-free [C-BSA][S₇]. The heating cycle of S₇ differs from that of solvent-free liquid [C-BSA][S₇] in that there is one exothermic glassy transition and there are two small endothermic glassy transitions that occur before the full melting transition. This indicates that there is an disordering and then a small reordering of surfactant molecules prior to melting, a behaviour that is removed once S₇ is incorporated into solvent-free liquid [C-BSA][S₇].

The presence of a full crystallization transition on the cooling cycle in both S₇ and solvent-free liquid [C-BSA][S₇] indicates that the kinetics of the system were faster than that of the scan rate, in contrast to S₁ and related solvent-free liquid proteins. The domain reorientation through the two glass transitions along with the temperature increase allows for the crystallization transition to occur within the heating cycle. The difference in thermal properties, *i.e.* the shift in melting and crystallization transition temperatures, between S₇ and solvent-free liquid [C-BSA][S₇] demonstrate that solvent-free

liquid [C-BSA][S₇] is a distinct material from S₇ formed of discrete particles, and not just BSA that has been dissolved within the S₇ surfactant.

3.2.3.1.3 DSC of solvent-free liquid [C-PK][S₇]

DSC traces of solvent-free liquid [C-PK][S₇] displayed a narrow reversible exothermic crystallization peak at 5 °C with a shoulder at 0 °C in the cooling cycle, and a sharp reversible endothermic melting transition peak at 45 °C in the heating cycle (Figure 3.24). These peaks were accompanied by an observable small glass transition occurring at -58 °C in the cooling cycle and a long but shallow glass transition occurring at -36 °C in the cooling cycle. The thermal transitions of solvent-free liquid [C-PK][S₇] again contrast the ones displayed on the DSC traces for S₇. This indicates that solvent-free liquid [C-PK][S₇] is a distinct material from S₇ formed of discrete particles, and not just PK that has been dissolved within S₇ surfactant. As with the solvent-free liquid proteins formed using S₁, the DSC traces of the two solvent-free liquid proteins differ, with the crystallization temperatures being either higher or lower than that of S₇, again demonstrating that the size of the protein being synthesised has an effect on the final materials thermal properties.

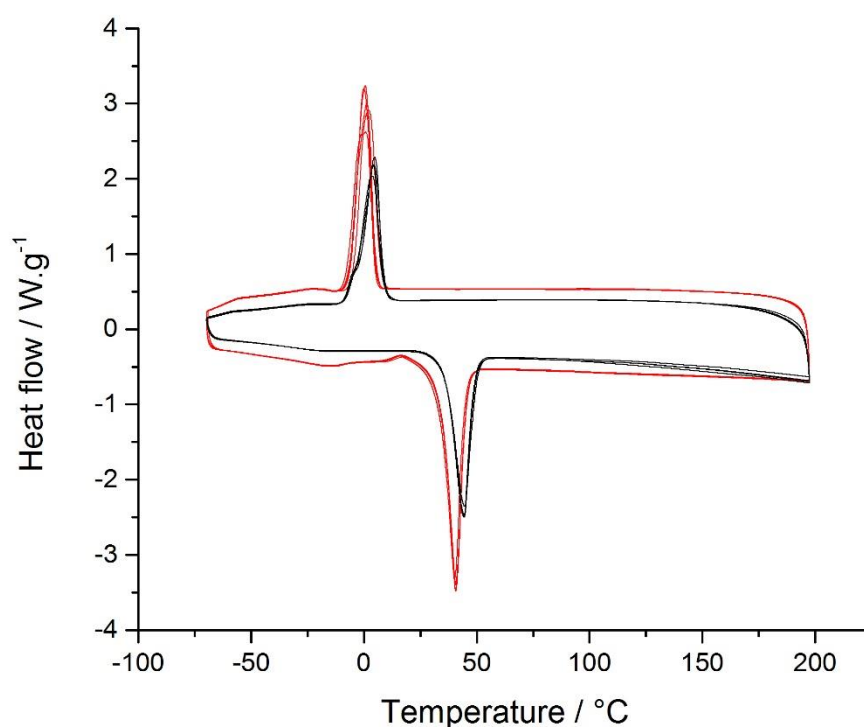


Figure 3.24: Cyclic DSC traces showing normalized heat flow against temperature for solvent-free liquid [C-PK][S₇] (black) and S₇ (red). [Experiments were cycled a minimum of 6 times and the scan rate was 10 °C.min⁻¹].

3.2.3.2 Thermogravimetric analysis

3.2.3.2.1 TGA of solvent-free liquid [C-BSA][S₁]

Thermo-gravimetric analysis (TGA) was used to analyse both the thermal stability and the final water content of the solvent-free liquid proteins. Each sample was held at 90 °C for 1 hour to remove any atmospheric water and then subsequently held at 110 °C to assess the water content of the material, with any change in weight relating to surface-bound water molecules. TGA of solvent-free liquid [C-BSA][S₁] showed a decomposition temperature of 402 °C and gave a residual water content of 0.05% (Figure 3.25). The thermal decomposition temperature of lyophilized BSA is 385 °C, displaying an increased thermal stability of the solvent-free liquid [C-BSA][S₁] construct of 17 °C. It is also important to note that the thermal decomposition temperature of S₁ was 336 °C, 62 °C lower than solvent-free liquid [C-BSA][S₁], again indicating that the material created is not just protein dissolved within surfactant, rather protein surfactant constructs with complex interactions.

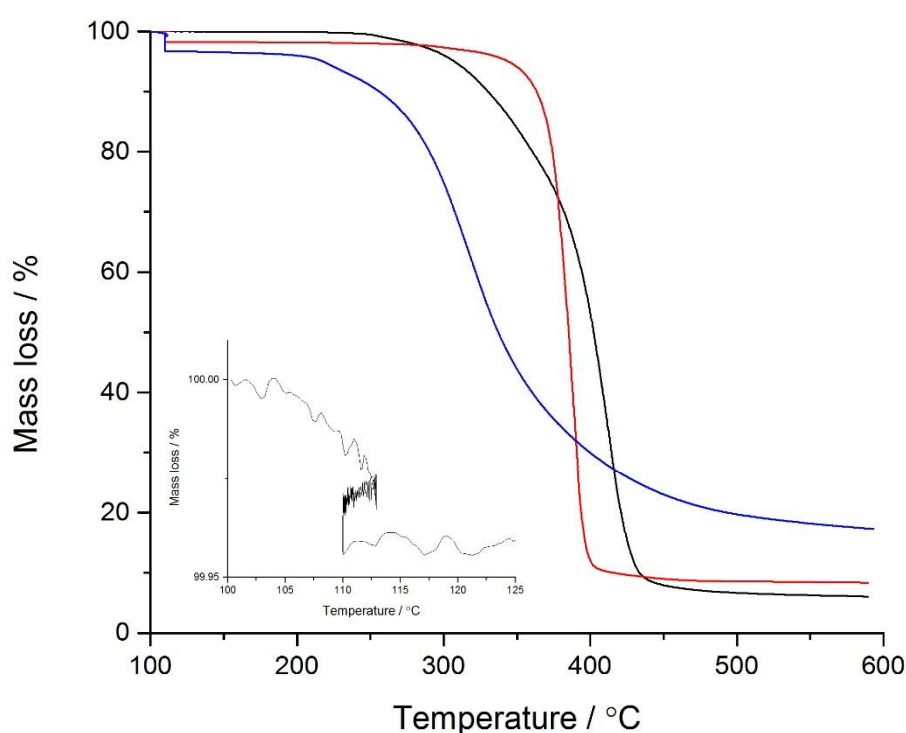


Figure 3.25: Thermogravimetric analysis of solvent-free liquid [C-BSA][S₁] (black), lyophilized BSA (blue), and S₁ surfactant (red), showing percentage mass loss in relation to temperature. The inset displays a magnified region of solvent-free liquid [C-BSA][S₁] in the range 100 - 120 °C showing the mass loss associated with the loss of surface-bound water (temperature held at 110 °C for 1 hour), *i.e.* the water content as a percentage.

For a 1:1 stoichiometric construct of solvent-free liquid [C-BSA][S₁], there are a total of 181 surfactant molecules present per protein. A 1:1 stoichiometry (where the ratio is the number of positively charged sites to the number of surfactant molecules) is assumed here due to previous work displaying

that when surfactant is added at a 2:1 ratio of number of positively charged sites to surfactant molecules, the final ratio after extensive dialysis is approximately 1:1¹⁰. There are 181 surfactant molecules present due to BSA having 181 positive surface charges that are solvent-accessible once cationized; 23 arginine, 40 aspartic acid, 59 glutamic acid, and 59 lysine³, with a cationization efficiency of 100% as calculated through mass spectrometry studies. The total mass of a solvent-free liquid [C-BSA][S₁] construct was calculated to be 298323 g.mol⁻¹; the mass of cationized BSA was 81666 g.mol⁻¹ and the mass of 181 S₁ surfactant molecules being 216657 g.mol⁻¹. With an experimentally derived water content of 0.05%, it can be calculated that each solvent-free liquid [C-BSA][S₁] construct has 8 surface-bound water molecules. The Pauling-Green theory²⁵ is a method used for the determination of the maximal water monolayer of a lyophilized protein, the theory calculates the number of tightly bound surface water molecules. The theory takes the number charged or highly polar residues in the primary sequence; aspartic acid, glutamic acid, serine, threonine, and tyrosine, and assumes one water molecule per site, with the 'mono-layer' covering distinct sites of the protein rather than a uniform layer surrounding the whole protein. These water molecules form hydrogen bonds with the sites and are essential for the correct folding of the protein and retention of the protein conformation. For native-BSA, the Pauling-Green theory predicts 181 charged or polar sites on the surface which translates to 181 water molecules in the water mono-layer. This yields a maximum mono-layer water content of 4.67%, 2 orders of magnitude higher than that calculated for each solvent-free liquid [C-BSA][S₁] construct. Analysis of how the loss of solvation shell effects the structure is detailed in Section 1.3.3.4.

With solvent-free liquid [C-BSA][S₁] having a cationization efficiency of 100% this indicates that all of the charged surface residues are positively charged and would, therefore, be interacting with surfactant molecules and not water molecules in a 1:1 stoichiometry. This means that these residues should be discounted when considering the Pauling-Green theory, leaving a potential mono-layer of 82 water molecules, relating to a water content of 2.2%. This water content value is still significantly higher than the 0.05%, or 8 water molecules, calculated for the solvent-free liquid [C-BSA][S₁] construct. The solvent accessible surface area (SASA) of a protein can be calculated from which the number of solvent molecules required to cover this area can be determined. For BSA the SASA is 67719 Å², requiring 11003 water molecules to fully cover, 4 orders of magnitude higher than the number present on the solvent-free liquid [C-BSA][S₁] construct. When determining where the highly associated water molecules are located on the solvent-free liquid [C-BSA][S₁] there are two main areas; the serine, threonine, and tyrosine residues, and the polyethylene glycol (PEG) unit in the surfactant chain. PEG is a hydrophilic molecule, strongly associating with water molecules through hydrogen bonds, with the water content of S₁ being 1.8% (Figure 3.25). The hydrophilic nature of PEG along with the high number of units present in the solvent-free liquid [C-BSA][S₁] (3620 PEG units) it stands to reason that a number of water molecules would associate with it, potentially reducing the number of water molecules in contact with the protein. This TGA data demonstrates that the solvent-

free liquid [C-BSA][S₁] construct is not solvated by a full water molecule hydration shell, and instead, only retains a vastly reduced number of strongly bound water molecules about the construct.

3.2.3.2.2 TGA of solvent-free liquid [C-BSA][S₇]

TGA of solvent-free liquid [C-BSA][S₇] displayed a decomposition temperature of 403 °C and a water content of 0.13% (Figure 3.26). Solvent-free liquid [C-BSA][S₇] therefore has a higher thermal stability compared to lyophilized BSA, an increase of 18 °C, a comparable increase to solvent-free liquid [C-BSA][S₁]. A water content of 0.13% was calculated for solvent-free liquid [C-BSA][S₇], again being similar to solvent-free liquid [C-BSA][S₁]; the total mass of a solvent-free liquid [C-BSA][S₇] construct was calculated to be 312260 g.mol⁻¹; the mass of cationized BSA was 81666 g.mol⁻¹ and the mass of 181 S₇ surfactant molecules being 230594 g.mol⁻¹, yielding 23 water molecules that are highly associated with the protein construct. Similarly to the solvent-free liquid [C-BSA][S₁] construct, the water content of a solvent-free liquid [C-BSA][S₇] construct was calculated to be lower than the predicted number of water molecules from the Pauling-Green theory and considerably lower than the number required to cover the SASA of the BSA protein. Again, these water molecules may not solely be associated with the protein, rather with the PEG units in the surfactant chains, S₇ having a water content of 5.5%. The increased number of PEG units in the solvent-free liquid [C-BSA][S₇] construct in relation to S₁ analogue (543 additional PEG units), may be a factor in the increased water content. This again confirms that the formation of a solvent-free liquid protein removes all solvent water molecules along with the majority of highly surface-associated water molecules from the protein.

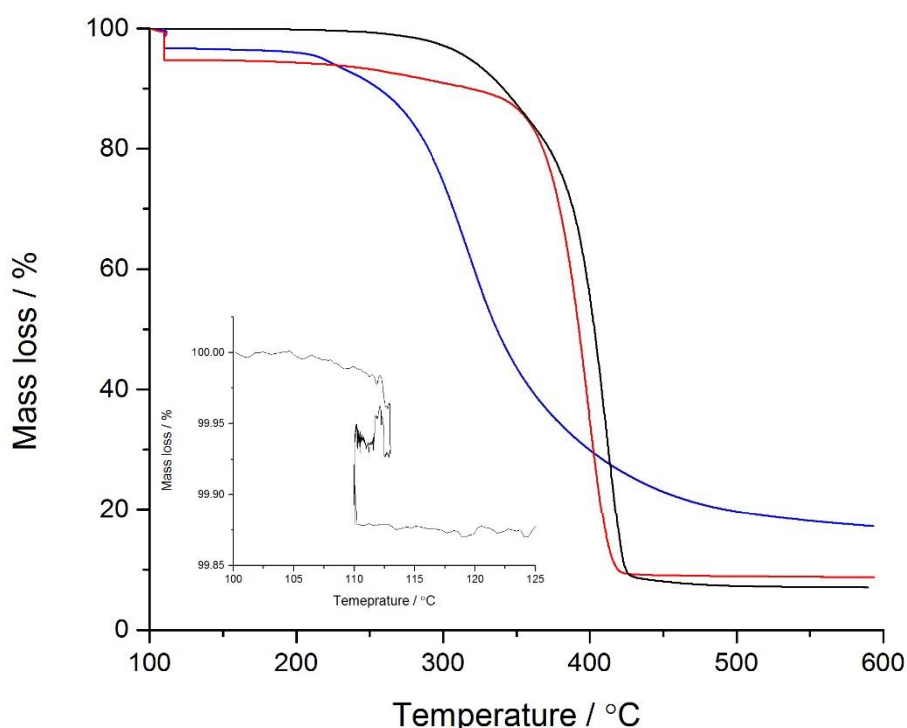


Figure 3.26: Thermogravimetric analysis of solvent-free liquid [C-BSA][S₇] (black), lyophilized BSA (blue), and S₇ surfactant (red), showing percentage mass loss in relation to temperature. The inset displays a zoomed-in region of solvent-free liquid [C-BSA][S₇] in the range 100 - 120 °C showing the mass loss associated with the loss of surface-bound water (temperature held at 110 °C for 1 hour), *i.e.* the water content as a percentage.

3.2.3.2.3 TGA of solvent-free liquid [C-PK][S₇]

TGA of solvent-free liquid [C-PK][S₇] displayed a decomposition temperature of 388 °C and a water content of 0.4% (Figure 3.27). The water content of 0.4 % for solvent-free liquid [C-PK][S₇] was also calculated to be comparable to solvent-free liquid [C-BSA][S₁]; the total mass of a solvent-free liquid [C-PK][S₇] construct was calculated to be 77794 g.mol⁻¹; the mass of cationized PK was 31930 g.mol⁻¹ and the mass of 38 S₇ surfactant molecules being 45864 g.mol⁻¹, yielding 6 water molecules that are highly associated with the protein construct. With solvent-free liquid [C-PK][S₇] having a cationization efficiency of 89%, 16 of the water molecules predicted from the Pauling-Green theory have to be discounted leaving a potential mono-layer of 78 water molecules out of, relating to a water content of 4.6%. This water content value is still significantly higher than the 0.4%, or 6 water molecules, calculated for the solvent-free liquid [C-PK][S₇] construct. The SASA of proteinase K was calculated to be 10809 Å², requiring 1756 water molecules to fully cover, 3 orders of magnitude higher than the number present on solvent-free liquid [C-PK][S₇] construct. It is important to note that it is not known whether the water molecules that are structurally significant to the active site have remained or removed through the formation of the solvent-free liquid. However, it is probable that they have remained over other surface bound water molecules due to the strong hydrogen bonding with multiple residues within the active site.

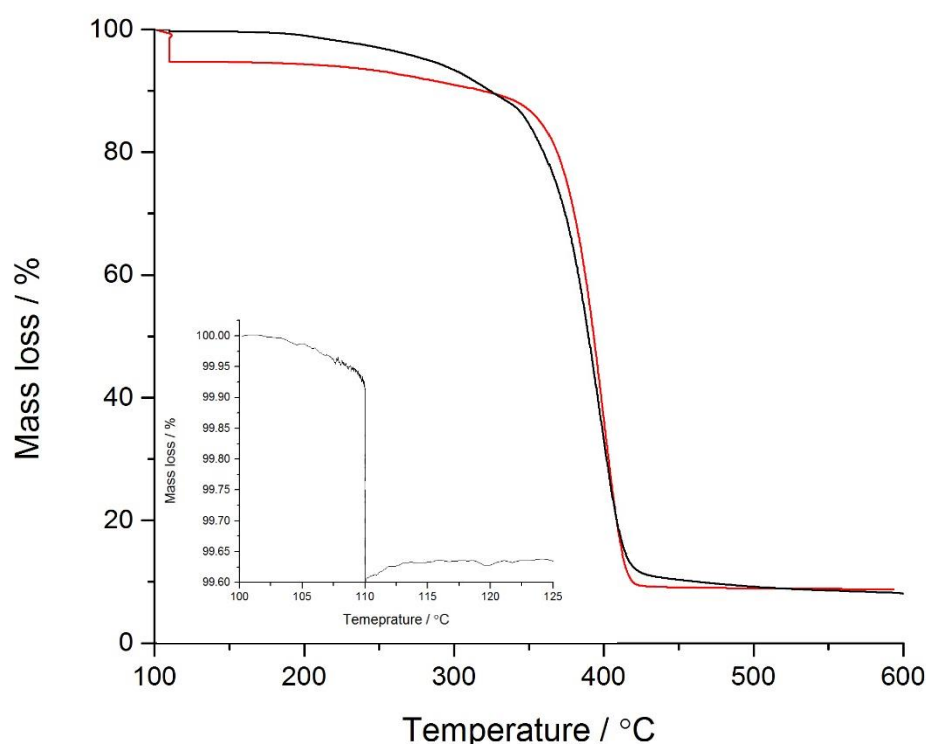


Figure 3.27: Thermogravimetric analysis of solvent-free liquid [C-PK][S₇] (black), and S₇ surfactant (red), showing percentage mass loss in relation to temperature. The inset displays a zoomed-in region of solvent-free liquid [C-PK][S₇] in the range 100 - 120 °C showing the mass loss associated with the loss of surface-bound water (temperature held at 110 °C for 1 hour), *i.e.* the water content as a percentage.

It should be noted that it is assumed that each construct is in a 1:1 ratio of positively charged site to surfactant, an assumption derived from previous work on solvent-free liquid proteins all showing no more than a ratio 1:1.3. This assumption is important as it changes the number of calculated water molecules associated with each surfactant-protein construct. However, even if the ratio was increased to 1:2, the maximum ratio due to the amount of surfactant added, the number of associated water molecules does not increase by a notable amount (Table 3.12).

Table 3.12: Table displaying the change in the number of water molecules associated with each solvent-free surfactant construct when the ratio of positively charged surface sites to the number of surfactant molecules is increased from 1:1 to 1:2

Sample	Ratio	Calculated number of water molecules
Solvent-free liquid [C-BSA][S ₁]	1:1	8
	1:2	14
Solvent-free liquid [C-BSA][S ₇]	1:1	9
	1:2	15
Solvent-free liquid [C-PK][S ₇]	1:1	2
	1:2	4

3.2.3.3 Comparing the thermal behaviours of differing surfactants and proteins

As previously discussed, the cause for the formation of a liquid protein phase after the addition of the surfactant corona is the extended inter-protein interactions. The head group, chain formation, and chain length (*i.e.* molecular weight) of the surfactant, along with the size of the protein and extent of cationization, will ultimately determine whether a solvent-free liquid protein can form and could also have an impact on how the solvent-free liquid protein behaves in relation to temperature. It is important to note that it is hard to distinguish which factors; the molecular weight, the head group, or the tail group, are causing the difference in behaviour in the solvent-free liquid proteins, as all are slightly different in S_1 and S_7 .

3.2.3.3.1 S_1 and S_7

Here, S_1 and S_7 both form solvent-free liquids of BSA, displaying that carboxylate and sulfonate head groups electrostatically interact with the cationized surface of the protein in a manner that is conducive to forming protein liquids. The decomposition temperatures for each protein liquid are approximately equal and higher than those of the lyophilized native BSA and of both respective individual surfactants, demonstrating an increase in thermal stability when the two individual components are associated. The TGA investigation displayed a 3-fold increase (0.05% - 0.13%) in water content for solvent-free liquid [C-BSA][S_7] over solvent-free liquid [C-BSA][S_1]. This is postulated to be caused by the slight increase in the number of PEG monomer units of S_7 (23) relative to S_1 (20), also seen in the water content of the surfactants, being 5.5 % and 1.1 % respectively. This minimal increase in water content is not thought to have had an effect on the thermal dynamics of the protein as the number of water molecules is still two orders of magnitude lower than that required for full solvation. A crucial difference between the BSA liquids formed with S_1 and S_7 appears in the DSC results, with the crystallization and melting temperature transitions differing. When BSA is used, the crystallization temperature increased from -5 °C to 2 °C, and the melting temperature decreased for S_1 relative to S_7 from 44 °C to 26 °C. This means that the energy required to form the solvent-free liquid state is lower in the case of S_1 , possibly indicating that large nonylphenyl domains in the tail of S_1 interact with lower energy than those in the straight alkyl chain of S_7 . This correlates well with the melting temperatures of S_1 and S_7 , at 32 and 40 °C respectively. As the crystallization temperature is lower for S_7 this means that the re-ordering of the straight alkyl chain domain takes longer and requires a lower energy to crystallize relative to the nonylphenyl domains. Interestingly this trend did not correlate well with the crystallization temperatures of S_1 and S_7 , -16 and 1 °C respectively, showing that not all intrinsic properties are carried through to the new material. This comparison indicates that the surfactant used has a large bearing on the thermal properties of the solvent-free liquid protein electrostatically bound to a cationized protein, with a straight chained surfactant having a higher melting point and a lower crystallization transition temperature relative to a surfactant with aromatic groups.

3.2.3.3.2 PK and BSA

Here, PK and BSA both formed solvent-free liquids, displaying that despite the larger molecular weight of BSA (66462 g.mol⁻¹) relative to PK (28993 g.mol⁻¹) and the lower cationization efficiency of CPK (89%) relative to CBSA (100%), it was possible to create solvent-free liquids of each protein. Solvent-free liquid [C-BSA][S₇] had a higher thermal decomposition temperature relative to solvent-free liquid [C-PK][S₇], at 403 and 388 °C respectively, indicating that solvent-free liquid [C-BSA][S₇] is more stable than solvent-free liquid [C-PK][S₇]. DSC of the two liquid proteins showed that the melting temperatures were approximately the same (\approx 45 °C), but the crystallization temperatures differed. Solvent-free liquid [C-BSA][S₇] crystallized at -5 °C, which was 10 °C lower than that of solvent-free liquid [C-PK][S₇] (5 °C), this was postulated to be an intrinsic property of the proteins since BSA is larger than PK it could take longer for it to re-fold and crystallize relative to PK. This comparison would therefore indicate that the intrinsic properties of the protein used has a significant impact on the thermal properties of the solvent-free liquid protein, with a larger more stable protein having a higher thermal resistance but also taking longer to re-crystallize.

3.2.3.4 Secondary structure analysis through circular dichroism spectroscopy (CD)

To interrogate the effect of the solvent-free liquid state on protein structure, CD spectroscopy experiments were conducted at 45 °C, to ensure that the solvent-free proteins were in the liquid state (45 °C is a higher temperature than the melt transition temperature for both solvent-free proteins). Each protein was held at this temperature for 5 minutes to ensure homogeneous melting throughout the sample.

3.2.3.4.1 CD of solvent-free liquid [C-BSA][S₇]

The formation of solvent-free liquid [C-BSA][S₇] increased the intensity of all characteristic peaks relative to C-BSA and aqueous [C-BSA][S₇], and also red-shifted the parallel $\pi \rightarrow \pi^*$ transition (208-211 nm), the $\pi \rightarrow \pi^*$ transition (193-195 nm), and the $n \rightarrow \pi^*$ transition (225-229 nm) (Figure 3.28). This returned the peaks to native-like wavelengths apart from the $n \rightarrow \pi^*$ transition which became shallow and more distorted relative to native-BSA (Figure 3.28). This exhibited greater structure reclamation than was seen through surfactant conjugation, with the polymer corona interactions creating a more native-like conformation in the modified state than those of water. The difference in the increase in intensity of the 208 nm peak relative to the 222 nm peak returned the $\theta_{222}/\theta_{205}$ ratio to \approx 0.63, similar to that of the C-BSA, again suggesting that 3_{10} helices are formed due to the de-stabilization of α -helices^{12,26,27}. However, the red-shifting of the 195 nm peak could suggest further reclamation of α -helical structure within the protein. The CD spectrum of solvent-free [C-BSA][S₇] implies potential structure reclamation, in the form of red-shifting characteristic peaks along with the increase in the intensity of these peaks relative to aqueous [C-BSA][S₇], whilst still remaining lower than native-BSA.

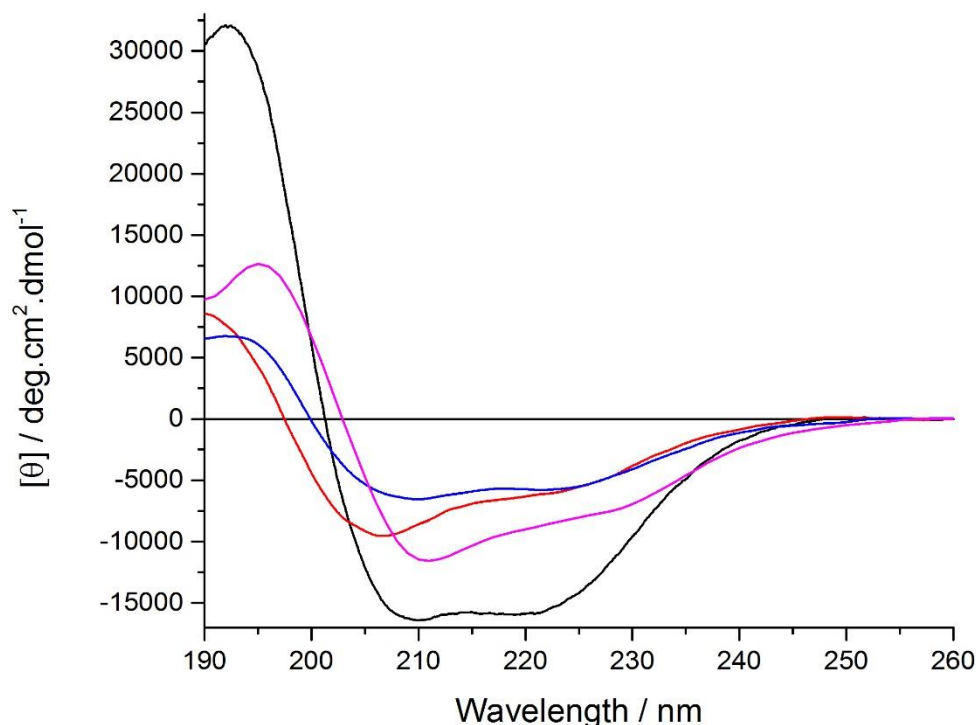


Figure 3.28: Mean residue ellipticity against wavelength showing far-UV CD spectra of aqueous BSA (black), C-BSA (red), and aqueous [C-BSA][S₇] (blue) at 45 °C, and far-UV SRCD spectrum of solvent-free liquid [C-BSA][S₇] (magenta) at 45 °C between 190-260nm.

Table 3.13: Table of secondary structure content calculated from CD spectra displayed in Figure 3.28.

Sample	α -Helix / %	β -Sheet / %	Turns / %	Other/ %
Solvent-free liquid [C-BSA][S ₇]	29	18	13	40
Aqueous [C-BSA][S ₇]	18	22	14	46
Aqueous C-BSA	18	26	14	42
Aqueous native-BSA	37	14	12	37

It is clear from the deconvoluted data that the creation of solvent-free [C-BSA][S₇] greatly altered the secondary structure of BSA relative to both aqueous [C-BSA][S₇] and C-BSA, with a large increase in α -helical content and a small decrease the β -sheet content; a change of 11% and 4% respectively (Table 3.13). This implies a shift that returns the secondary structure to a more native-like conformation, a trend that has been observed in previous solvent-free liquid protein work¹⁰. The cause of this shift towards a more highly folded, and therefore more densely packed, protein structure has previously been rationalised by a decrease in the entropy of unfolding once in the solvent-free liquid state. Even though the enthalpy of the unfolding transition correspondingly increases, there remains a disparity in free energy. This arises from a molecular crowding effect brought on through the reduction in conformational freedom once in the solvent-free liquid state, a state that densely packs the interior

residues of globular proteins²⁸. Therefore, the replacement of all solvation-water/protein hydrogen bonding interactions with surfactant/protein interactions increased the secondary structure of the modified-BSA relative to the aqueously dispersed construct, with the modified-BSA displaying a high retention of secondary structure content relative to native-BSA.

3.2.3.4.2 CD of solvent-free liquid [C-PK][S₇]

The formation of solvent-free [C-PK][S₇] reduced the intensity of two characteristic peaks (208 and 222 nm), induced a narrowing of the 222 and 208 nm peaks, along with a marginal increase in the $\theta_{222}/\theta_{208}$ ratio (Figure 3.29). As well as the narrowing of the 222 and 208 nm peaks there is a blue shift of the $n \rightarrow \pi^*$ transition peak (223-222 nm), of the positive $\pi \rightarrow \pi^*$ transition peak (190 - 194 nm), and of the parallel $\pi \rightarrow \pi^*$ transition peak (207-210 nm). This returns the peak positions to similar wavelengths to those of native-PK, with the decrease in signal intensity of the 222 and 208 nm characteristic peaks retaining a similar $\theta_{222}/\theta_{208}$ ratio relative to aqueous [C-PK][S₇] (1.38) (remaining at a higher value than that of native-PK). The CD spectrum of solvent-free [C-PK][S₇] displays potential structure reclamation, in the form of red-shifting characteristic peaks to wavelengths similar to that of native-BSA, with the CD spectrum approximately overlaying the native-BSA CD spectrum.

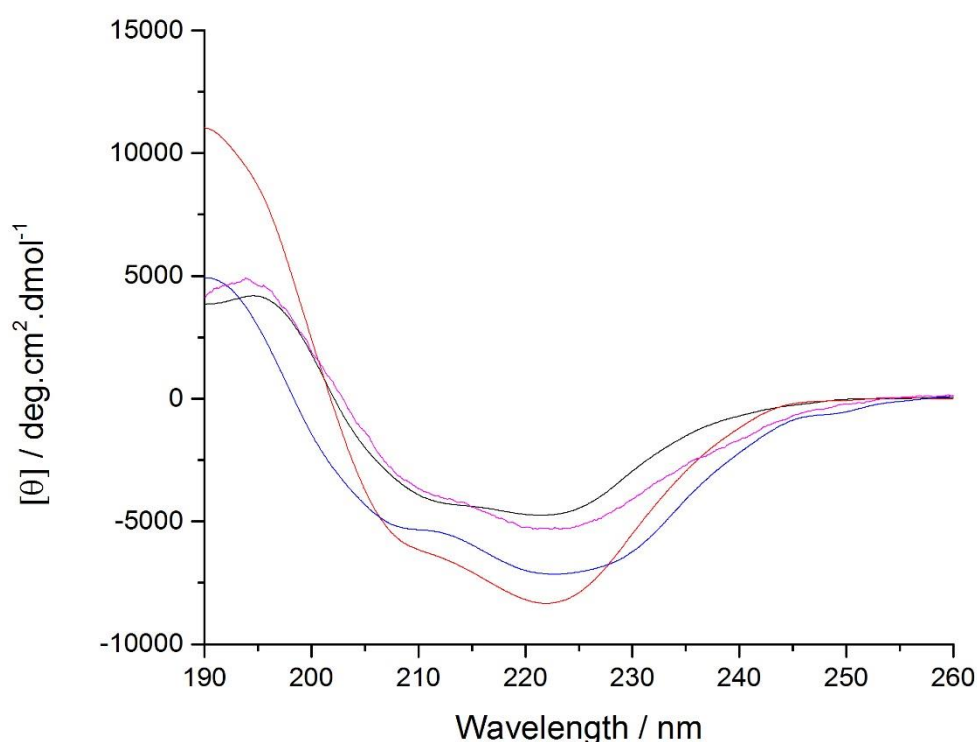


Figure 3.29: Mean residue ellipticity against wavelength showing far-UV CD spectra of aqueous PK (black), C-PK (red), and aqueous [C-PK][S₇] (blue) at 45 °C, and far-UV SRCD spectrum of solvent-free liquid [C-PK][S₇] (magenta) at 45 °C between 190-260nm.

It is clear from the deconvoluted data that the creation of solvent-free [C-PK][S₇] did not alter the secondary structure of the protein substantially, relative to aqueous [C-PK][S₇] (Table 3.144). The removal of all solvation water molecules only appeared to decrease the α -helical content marginally and increase the β -sheet content by 5%. The deconvoluted data fits with the observations in the CD spectra, the increase in β -sheet content causing the marginal increase in the $\theta_{222}/\theta_{208}$ ratio. The retention of structure relative to the aqueously dispersed [C-PK][S₇] conjugate and native-PK shows that the presence of polymer surfactant both screens the additional repulsive surface charges and approximately replicates the interaction environment given to the protein by the presence of solvation water molecules. However, as mentioned previously, the marginal alteration in structure content observed could be detrimental to the activity of the protease.

Table 3.14: Table of secondary structure content calculated from CD spectra displayed in Figure 3.29.

Sample	α -Helix / %	β -Sheet / %	Turns / %	Other / %
Solvent-free liquid [C-PK][S ₇]	9	31	15	43
Aqueous [C-PK][S ₇]	10	26	17	46
Aqueous C-PK	18	33	13	37
Aqueous native-PK	10	34	13	43

3.2.3.5 Thermal denaturation of solvent-free liquid proteins

Thermal denaturation experiments were again conducted to investigate the unfolding mechanics and thermal stability of the modified proteins. Since solvent-free liquid proteins do not contain water, the inherent testing temperature restriction for native protein of 95 °C, due to its aqueous environment, no longer applies, thus allowing for higher temperature thermal analysis. However, there is an additional testing temperature restriction for solvent-free proteins due to their solid state at temperatures below 40 °C, the analysis therefore cannot begin until the protein melts. The transition to the liquid phase causes a large structural rearrangement of the protein (Figure 3.30) which made using the selected model for calculation of thermodynamic values impossible. This dislocation in data is due to the crystallised solvent-free modified protein melting and rearranging the structure to that of the solvent-free liquid protein. To ensure that homogenous melting occurred throughout, each protein sample was held at 45 °C for 5 minutes prior to taking any measurements.

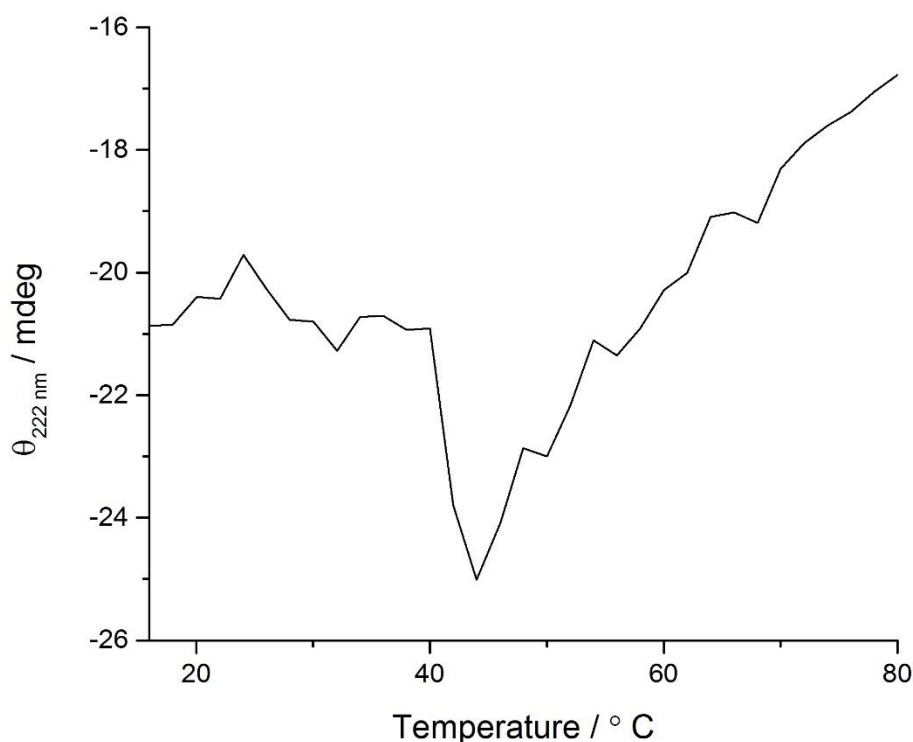


Figure 3.30: Ellipticity at 222 nm against temperature showing the far-UV SRCD 222nm spectrum of solvent-free liquid [C-BSA][S₇] between a temperature range of 16-80 °C.

3.2.3.5.1 Thermal denaturation of solvent-free liquid [C-BSA][S₇]

The creation of the solvent-free liquid state of [C-BSA][S₇] removed the characteristic sigmoidal thermal unfolding curve as displayed in native-BSA, with the protein gradually unfolding as temperature increases (Figure 3.35B). This behaviour continues linearly until approximately 80 °C where the rate of unfolding decreases. Eventually the structure became fully denatured at approximately 140 °C, concurring with similar full denaturation temperature in previous literature on BSA²⁹. Due to the denaturation curve not being sigmoidal in nature, thermodynamic values were not calculable using the selected model.

A thermal denaturation curve of this nature suggests that whilst the replacement of solvation water molecule interactions with polymer interactions reclaims a more native-like secondary structure relative to C-BSA and aqueous [C-BSA][S₇], the protein becomes molten-globular in nature. This effect has been observed previously with other solvent-free liquid proteins¹⁰. This behaviour indicates that the polymer-protein interactions do not offer the same interaction environment relative to water-protein interactions; the protein being in a more loosely bound native-like conformation (molten-globular like). The readiness of the protein to unfold in the polymer-rich environment could be related to the increase in the hydrophobic nature of the protein's surroundings caused by the presence of a hydrophobic section of the surfactant tail. This increase in hydrophobicity reduces the enthalpy barrier

to unfolding as it is now more energetically favourable for the hydrophobic core-residues to interact with these sections of the surfactant. This effect was also observed previously in other solvent-free liquid protein studies⁷. The removal of the structural dependency on water molecules does allow the structure of BSA to be examined past the boiling point of water. This means the formation of solvent-free liquid [C-BSA][S₇] provides an alternative approach for the full analysis of the unfolding mechanics of BSA, permitting the fully denatured form to be present.

It is clear from the de-convoluted unfolding structure data that as solvent-free liquid [C-BSA][S₇] unfolds the α -helices present rapidly convert to β -sheet structures, the cross-over temperature of dominant secondary structure percentage being at 80 °C (Figure 3.31C). The presence of large percentages of secondary structure motifs at these higher temperatures does display an increase in thermal stability relative to aqueous [C-BSA][S₇]. The α -helical content is 0% at 80 °C for aqueous [C-BSA][S₇] (a drop of 100%), whereas it is approximately 20% for solvent-free liquid [C-BSA][S₇] at the same temperature (a drop of approximately 30%). This high retention of the structure at elevated temperatures has been previously explained in terms of the entropy of denaturation. In the highly concentrated solvent-free liquid form the proteins are crowded close together, such that it is entropically unfavourable for them to unfold, increasing the energy required for this to occur⁶. Therefore, whilst the enthalpy barrier to denaturation has decreased, due to the increase in hydrophobicity, the entropic barrier to denaturation has increased due to the large increase in concentration causing molecular crowding.

Therefore, whilst the unfolding mechanism of solvent-free liquid [C-BSA][S₇] differs from that of native-BSA due to the formation of a molten-globular state, with an increase in thermal stability arising due to entropic trapping of the protein in this state, retaining a higher percentage of secondary structure at an elevated temperature. This, therefore, provides an alternative method for analysing the structure of BSA past the solvation-based temperature barrier, allowing for analysis of a fully denatured native-like BSA. It is assumed that due to the comparable chain structure, length, and hydrophobic/hydrophilic characteristics of S1 relative to S7, that solvent-free liquid [C-BSA][S₁] is structurally commensurate and therefore thermodynamically similar to solvent-free liquid [C-BSA][S₇].

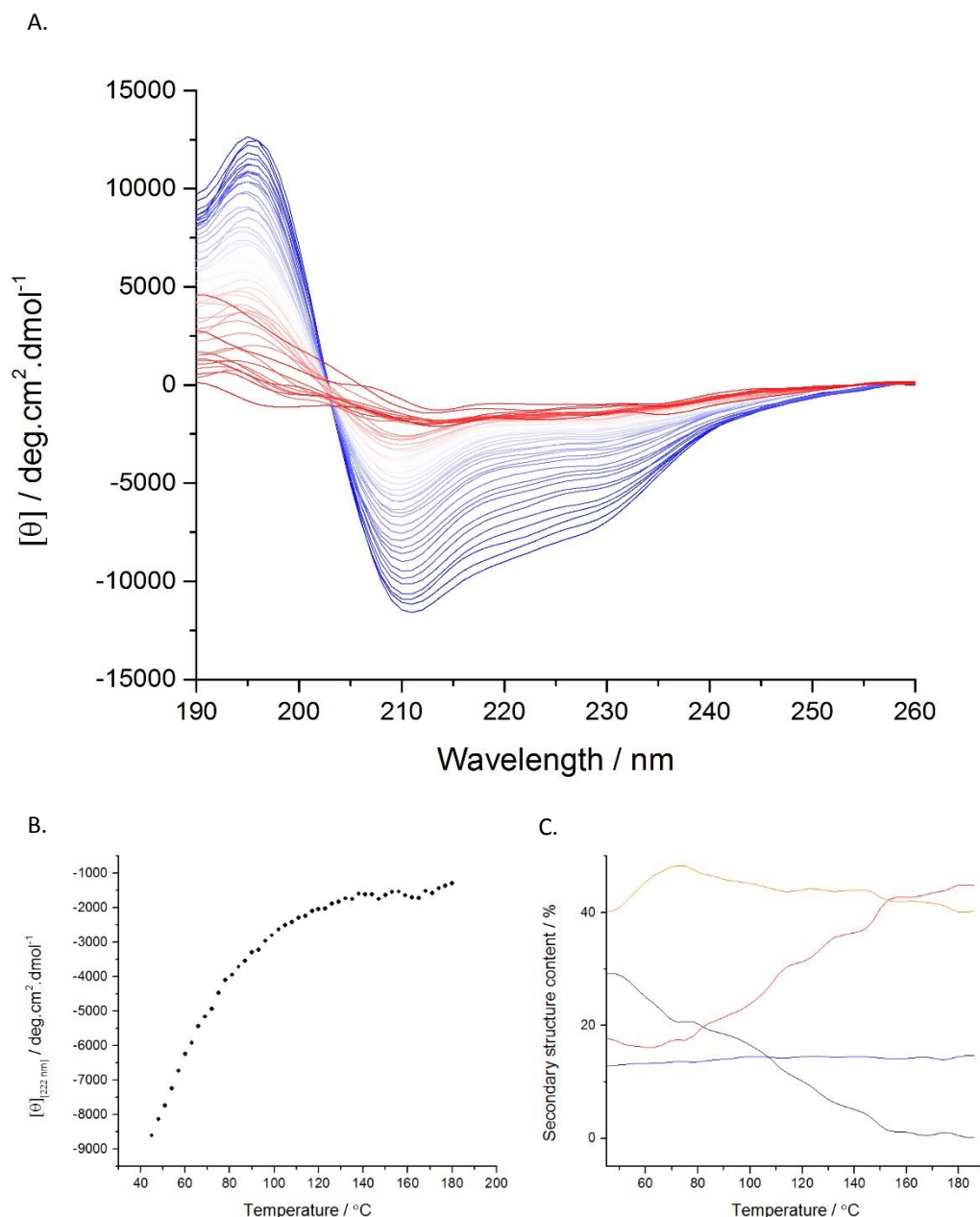


Figure 3.31: **A.** Mean residue ellipticity against wavelength showing far-UV SRCD spectra for solvent-free liquid [C-BSA][S₇], with the gradient from blue to red indicating an increase in temperature from 45 – 186 °C, **B.** Mean residue ellipticity at 222 nm against temperature (45 – 186 °C) for solvent-free liquid [C-BSA][S₇], and **C.** secondary structure percentage change against temperature (45 – 186 °C) for solvent-free liquid [C-BSA][S₇] (secondary structure percentage calculation detailed in Chapter 2), with the percentage of α-helices (black), β-sheets (red), random coil (orange), and turns (blue).

3.2.3.5.2 Thermal denaturation of solvent-free liquid [C-PK][S₇]

The creation of the solvent-free liquid state of [C-PK][S₇] removed the characteristic double sigmoidal thermal denaturation curve as displayed by all three aqueous precursors (Figure 3.32B). The denaturation curve instead displayed an unexpected unfolding and re-folding event, with the potential formation of a stable native-like intermediate species at approximately 80 °C that remained stable until approximately 130 °C (Figure 3.32). Due to the denaturation curve not being sigmoidal in nature, thermodynamic values were not calculable using the selected model.

The unfolding curve of solvent-free liquid [C-PK][S₇] did not match the native protein unfolding curve, or that of the cationized or surfactant conjugated species either, indicating a deviation in thermal denaturation mechanism. The initial unfolding occurs linearly until 58 °C, where the protein begins to re-fold at a marginally slower rate, stabilizing at 85 °C. This indicates that the initial conformation of the protein is not stable at elevated temperatures, instead favouring a similarly folded native-like intermediate conformation. The potential intermediate species remains stable until 128 °C, where the protein again began to unfold. The initial unfolding and re-folding events are all absent from the deconvoluted secondary structure, with only the second unfolding event being displayed as a marginal increase in β -sheet content and a decrease in α -helical content (Figure 3.32C). The lack of change in deconvoluted data despite clear changes in the CD spectra is an apparent trend with small changes in signal intensity with all proteinase K species. As this thermal denaturing mechanism is vastly different from aqueous [C-PK][S₇], the only cause for this deviation can be attributed the removal of water molecules from the structure. It is likely, as with solvent-free liquid [C-BSA][S₇], that this modification leads to a large increase in molecular crowding bringing about an increase in the entropic barrier to denaturation. This is despite the decrease in the enthalpic barrier brought about by the increase of hydrophobicity lowering the energetic barrier for the hydrophobic core residues to be displayed. This entropic barrier permitted the retention of a large percentage of the secondary structure, approximately 90% remaining up to 140 °C.

Therefore, whilst the unfolding mechanism of solvent-free liquid [C-PK][S₇] differs from that of the aqueous precursors, there is a large increase in the thermal stability of PK, retaining a higher percentage of the native-like secondary structure at an elevated temperature. As the structure is potentially highly native-like at these elevated temperatures it is likely that high temperature proteolytic activity could be observed. However, as detailed previously, any of the surface modifications conducted could render the protease in-active.

It is assumed that due to the comparable chain structure, length, and hydrophobic/hydrophilic characteristics of S₁ relative to S₇, that solvent-free liquid [C-PK][S₁] is structurally commensurate and therefore thermodynamically similar to solvent-free liquid [C-PK][S₇].

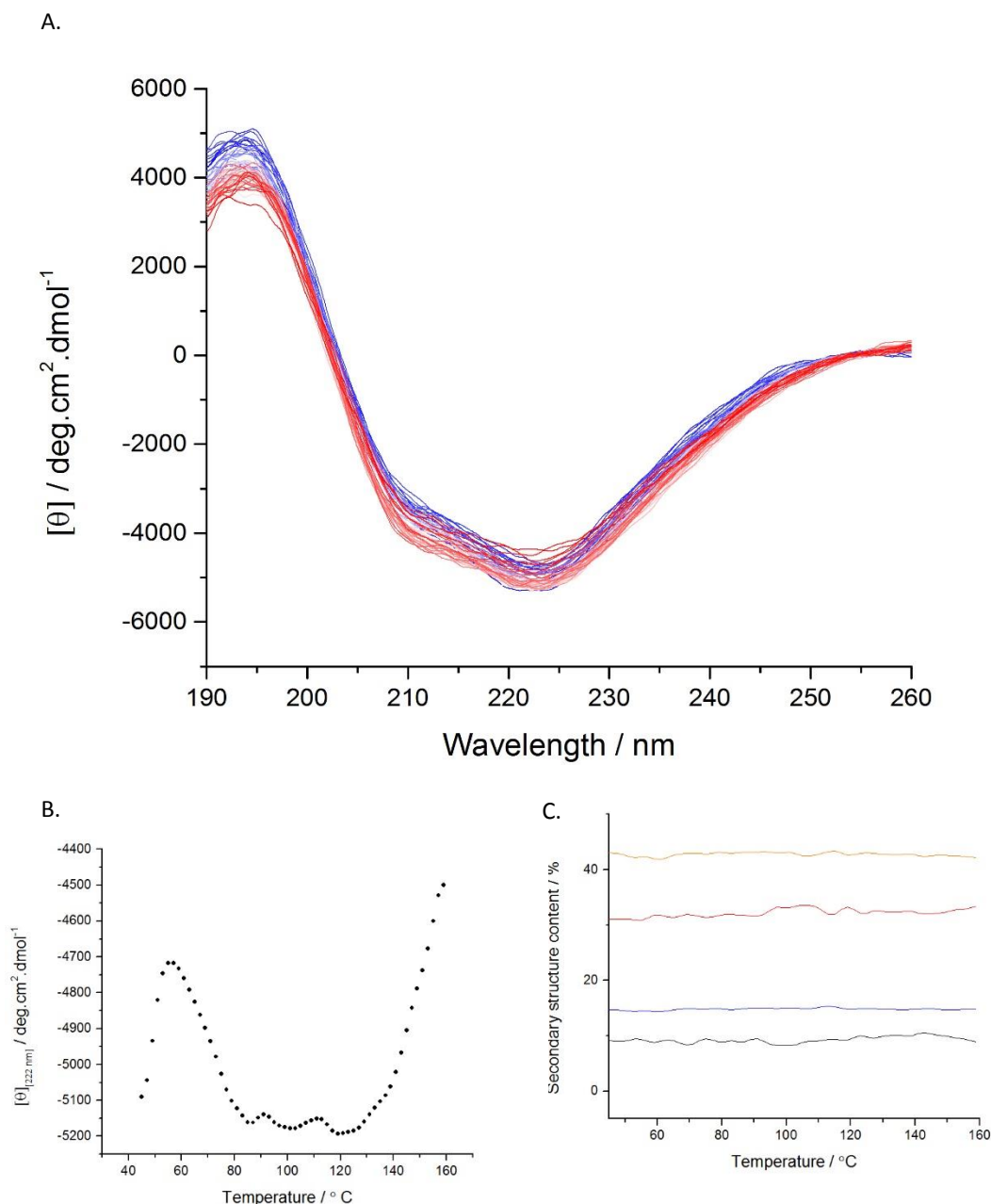


Figure 3.32: **A.** Mean residue ellipticity against wavelength showing far-UV SRCD spectra for solvent-free liquid [C-PK][S₇], with the gradient from blue to red indicating an increase in temperature from 45 – 160 °C, **B.** Mean residue ellipticity at 222 nm against temperature (45 – 160 °C) for solvent-free liquid [C-PK][S₇], and **C.** secondary structure percentage change against temperature (45 – 160 °C) for solvent-free liquid [C-PK][S₇] (secondary structure percentage calculation detailed in Chapter 2), with the percentage of α -helices (black), β -sheets (red), random coil (orange), and turns (blue).

3.2.3.6 *Conclusions for solvent-free liquid proteins*

The synthesis of four novel solvent-free proteins was successful; [C-BSA][S₁], [C-BSA][S₇], [C-BSA][S₁], and [C-PK][S₇]. The characterisation of each solvent-free liquid protein yielded interesting results that were generally comparable to previous solvent-free liquid proteins in the literature.

Differential scanning calorimetry revealed that the melting and recrystallization transitions of the solvent-free liquid proteins were different to that of the corresponding surfactants, providing evidence towards the material being a surfactant nanoconjugate and not a mixture of non-interacting surfactant polymer and protein. Thermogravimetric analysis of the solvent-free liquid proteins also provided supporting evidence towards this, with the thermal decomposition of the solvent-free liquid proteins differing from those of the native lyophilized proteins and pure surfactant. This technique also yielded a value for the number of water molecules per protein construct, in all four cases the value being lower than 30, at least 2 or 3 orders of magnitude lower than that required for minimal solvation of the native proteins.

It is interesting to note that the creation of a solvent-free liquid form of both proteins had different effects regarding structure; solvent-free liquid [C-BSA][S₇] became more native-like, whilst solvent-free [C-PK][S₇] remained approximately the same as aqueous [C-PK][S₇]. This further exemplifies the understanding that each protein, and therefore the surface environment of each protein, is different. Each modification would therefore alter the secondary and tertiary structure, the surface charge pairs, and thermal dynamics to varying positive and negative extents, all depending on the native protein. However, it is evident from the work presented here, along with previous solvent-free liquid protein investigations, that the replacement of water/protein hydrogen bonding interactions with polymer/protein interactions is not detrimental to the structure of the protein. Instead, it is evident that there is a regaining of structure lost through the various surface modifications employed. This reclamation of structure could indicate a similar reclamation in protein activity, however, an activity study would be required to show this.

The thermal denaturation of the solvent-free liquid proteins exhibited different unfolding mechanisms both to each other and their respective aqueous precursors. Solvent-free liquid [C-BSA][S₇] unfolded in a manner indicating the creation of a molten-globular form of the protein; gradually denaturing rather than taking a sigmoidal denaturation route. Interestingly, solvent-free liquid [C-PK][S₇] displayed the potential formation of a native-like intermediate species that remained stable up to 128 °C, differing from the double sigmoidal characteristic displayed by all three aqueous precursors. This difference in structural dynamics in response to thermal change is possibly due to the original change in secondary structure once cationized, with PK being more resistant to structure alteration once cationized and retaining the formation of an intermediate species. As the CSA was already in a higher state of disorder and did not refold towards the native-like structure to a large degree once conjugated

with surfactant, then it follows that it would not unfold with native-like characteristics. However, both solvent-free liquids displayed the retention of secondary structures over 100 °C, with PK retaining a native-like conformation to over 140 °C. This agrees with the previously documented thermal stability increase brought about through the formation of other solvent-free liquid proteins, thought to be caused by the entropic entrapment of the protein through a molecular crowding effect²⁸.

3.3 Overall conclusions and future work

3.3.1 Conclusions

The effect that the solvent-free liquid protein synthesis had on the two proteins utilized, *bovine serum albumin* (BSA) and proteinase K (PK), was investigated. The work was conducted primarily using CD spectroscopy, investigating the effect to secondary structure and thermal denaturation relative to the native species of each protein. Each surface modification utilized in the solvent-free liquid protein synthesis was shown to have varying effects on the proteins, as detailed previously and summarised below.

The optimized cationization of BSA and PK through the coupling of amides to negatively charged surface sites increased the number of positively charged surface sites present. This had the effect of removing surface ion pairings and salt bridges on both proteins, replacing them with repulsive pairings. Altering the electrostatic surface environment changed the structures and thermal stability of both proteins. C-BSA lost a large percentage of the secondary structure relative to native-BSA, with the presence of aggregates in the system indicating that this denaturation exposed hydrophobic core residues leading to aggregation stabilized through the formation of inter-protein β -sheets. A molten-globular state was also formed, signifying that the ion pairs that were disrupted were key to the stabilization of the native-BSA conformation. Contrastingly, C-PK gained secondary structure content relative to native-PK and did not form aggregates once cationized, displaying how the different surface environments response to the same process. The thermal denaturation pathway of C-PK also differed relative to native-PK, displaying similar transition temperatures but with the stabilization of a different intermediate species.

The 1:1 electrostatic surfactant addition to each positively charged surface site of both cationized proteins had the effect of shielding the repulsive electrostatic surface pairs, again changing the structure and thermal stability of both proteins. Aqueous [C-BSA][S₇] nanoconjugate displayed the dissolution of the majority of the C-BSA aggregates with a congruent marginal reclamation of native secondary structure, in the form of a reduction in β -sheet content, indicating that the dispersal of the C-BSA aggregates was due to re-burial of hydrophobic core residues removing the need for aggregate stabilization. A more globular-like state was regained, with some reclamation of a sigmoidal thermal denaturation, although the half denaturation temperature was lower relative to native-BSA. This indicates that, whilst there was structure reclamation, the reclaimed intramolecular interactions are not of the same energy as native-BSA. Also, the entropy and enthalpy of denaturation were approximately 1/3rd relative to the native, indicating that the structure is less thermally stable. Aqueous [C-PK][S₇] displayed reclamation of secondary structure to near native-like content, along with retention of the thermal denaturation pathway. Whilst the unfolding mechanism was similar, it

was evident that the intermediate species stabilized was again different relative to native-PK and C-PK.

The removal of all solvation water molecules from the surface of the proteins through the formation of a solvent-free liquid protein, again, had differing effects on the structure and thermal stability of both proteins. This variation is likely rooted in how the previous surface modifications changed the structures of the two different proteins, with some alterations potentially not being reversible or reclaimable by the formation of the solvent-free liquid state. Also, the aqueous [C-PK][S₇] nanoconjugate was already in a near-native-like conformation, whilst the aqueous [C-BSA][S₇] nanoconjugate was in a molten globular like state, indicating that the previous modifications had had a more negative impact on the BSA.

The formation of solvent-free liquid [C-BSA][S₇] displayed a large reclamation in secondary structure content, returning the protein to a more native-like conformation. The thermal degradation pathway resembled that of a molten-globular protein, gradually unfolding until fully denatured. However, there was a large percentage of secondary structure present at higher temperatures than that of native-BSA indicating that solvent-free liquid [C-BSA][S₇] is more thermally stable than native-BSA. Solvent-free liquid [C-PK][S₇] displayed a marginal alteration to the secondary structure, bringing it closer in-line with native-PK. The native-like thermal degradation pathway was completely lost, with the protein unfolding and re-folding to form a native-like intermediate that remained stable until 128 °C, where it began to unfold. As solvent-free liquid [C-PK][S₇] displayed a high percentage of secondary structure at higher temperatures relative to native-PK it is more thermally stable. The large increase in thermal stability exhibited by both solvent-free liquid proteins is explained by the reduction in conformational freedom caused by the very high concentration of protein within the liquid.

The analysis of solvent-free liquid [C-BSA][S_x] and [C-PK][S_x] by differential scanning calorimetry and thermogravimetric analysis showed evidence that the liquids created were formed through the interactions between individual surfactant-nanoconjugates and not just a mixture of protein and surfactant. This was justified by the melting transition- and thermal decomposition- temperatures of the solvent-free liquids differing from both the pure surfactants and the respective native proteins. Both solvent-free liquid proteins were also shown to retain only 10s of water molecules, 2-3 magnitudes lower than that required for solvation. Interestingly, birefringence is displayed in both crystallized solvent-free [C-BSA][S] and [C-PK][S], unlike solvent-free liquid [C-Myoglobin][S]. This indicates that the optical effect may be dependent on the size of the protein (both BSA and PK having larger radii of hydration than Myoglobin).

Overall, it has been shown that the solvent-free liquid protein synthesis yields solvent-less native-like protein liquids that have a higher thermal stability relative to the respective native proteins. The polymer-surfactant corona that is formed through surface modifications replaces the hydration shell

around the protein, increasing the interaction distance between the proteins, permitting the formation of a liquid protein phase that reversibly melts and solidifies with minimal hysteresis.

3.3.2 Future work

To further the analyse into how each modification alters the BSA and PK, other techniques that have been used in previous solvent-free liquid protein characterisations could be employed, such as; Fourier transform infra-red spectroscopy (FT-IR), giving insight into how the amide bond environments change in these two proteins⁷; rheometry, exploring further the rheological differences that using different sized proteins induces⁹; and small-angle neutron scattering²³, shedding light on the structure of the solvent-free liquid proteins and their aqueous precursors. Each technique would further the knowledge of these unique materials and allow for intelligent design of future functional solvent-free protein materials.

3.4 References

1. Perriman, A. W. *et al.* Reversible dioxygen binding in solvent-free liquid myoglobin. *Nat. Chem.* **2**, 622–6 (2010).
2. McKenzie, H. A., Smith, M. B. & Wake, R. G. Molecular weight of ovalbumin and of bovine serum albumin in urea solution. *Nature* **176**, 738 (1955).
3. Majorek, K. A. *et al.* Structural and immunologic characterization of bovine, horse, and rabbit serum albumins. *Mol. Immunol.* **52**, 174–182 (2012).
4. Sweeney, P. J. & Walker, J. M. in *Enzymes of Molecular Biology* (ed. Burrell, M. M.) 305–311 (Humana Press, 1993). doi:10.1385/0-89603-234-5:305
5. Wang, J., Dauter, M. & Dauter, Z. What can be done with a good crystal and an accurate beamline? *Acta Crystallogr., Sect. D* **62**, 1475–1483 (2006).
6. Brogan, A. P. S., Sharma, K. P., Perriman, A. W. & Mann, S. Enzyme activity in liquid lipase melts as a step towards solvent-free biology at 150 °C. *Nat. Commun.* **5**, 6058 (2014).
7. Brogan, A. P. S., Sharma, K. P., Perriman, A. W. & Mann, S. Isolation of a highly reactive β -sheet-rich intermediate of lysozyme in a solvent-free liquid phase. *J. Phys. Chem. B* **117**, 8400–7 (2013).
8. Nagy, B. *et al.* Small angle neutron scattering study of globular proteins confined in porous carbons. *Carbon N. Y.* **106**, 142–151 (2016).
9. Perriman, A. W., Cölfen, H., Hughes, R. W., Barrie, C. L. & Mann, S. Solvent-free protein liquids and liquid crystals. *Angew. Chemie - Int. Ed.* **48**, 6242–6246 (2009).
10. Brogan, A. P. S., Siligardi, G., Hussain, R., Perriman, A. W. & Mann, S. Hyper-thermal stability and unprecedented re-folding of solvent-free liquid myoglobin. *Chem. Sci.* **3**, 1839 (2012).
11. Yazawa, K., Sugahara, M., Yutani, K., Takehira, M. & Numata, K. Derivatization of Proteinase K with Heavy Atoms Enhances Its Thermal Stability. *ACS Catal.* **6**, 3036–3046 (2016).
12. Andersen, N. H., Liu, Z. & Prickett, K. S. Efforts toward deriving the CD spectrum of a 3 10 helix in aqueous medium. *FEBS Lett.* **399**, 47–52 (1996).
13. Abidi, M., Iram, A., Furkan, M. & Naeem, A. Secondary structural alterations in glucoamylase as an influence of protein aggregation. *Int. J. Biol. Macromol.* **98**, 459–468 (2017).
14. Bhattacharya, M., Jain, N. & Mukhopadhyay, S. Insights into the mechanism of aggregation and fibril formation from bovine serum albumin. *J. Phys. Chem. B* **115**, 4195–4205 (2011).
15. Fink, A. L. Protein aggregation: Folding aggregates, inclusion bodies and amyloid. *Fold. Des.* **3**, 9–23 (1998).
16. Rambaran, R. N. & Serpell, L. C. Amyloid fibrils. *Prion* **2**, 112–117 (2008).
17. Sethuraman, A. & Belfort, G. Protein structural perturbation and aggregation on homogeneous surfaces. *Biophys. J.* **88**, 1322–1333 (2005).
18. Choy, N., Raussens, V. & Narayanaswami, V. Inter-molecular coiled-coil formation in human apolipoprotein E C-terminal domain. *J. Mol. Biol.* **334**, 527–539 (2003).
19. Kiss, R. S. *et al.* Structure-guided protein engineering modulates helix bundle exchangeable apolipoprotein properties. *J. Biol. Chem.* **278**, 21952–21959 (2003).
20. Berezovsky, I. N., Zeldovich, K. B. & Shakhnovich, E. I. Positive and negative design in stability and thermal adaptation of natural proteins. *PLoS Comput. Biol.* **3**, 0498–0507 (2007).
21. Muhaxhiri, Z. *et al.* Structural Basis of Substrate Specificity and Protease Inhibition in Norwalk

-
- Virus. *J. Virol.* **87**, 4281–4292 (2013).
22. Sharma, K. P., Collins, A. M., Perriman, A. W. & Mann, S. Enzymatically active self-standing protein-polymer surfactant films prepared by hierarchical self-assembly. *Adv. Mater.* **25**, 2005–10 (2013).
 23. Brogan, A. P. S., Sessions, R. B., Perriman, A. W. & Mann, S. Molecular Dynamics Simulations Reveal a Dielectric-Responsive Coronal Structure in Protein-Polymer Surfactant Hybrid Nanoconstructs. (2014).
 24. Gallat, F.-X. *et al.* A polymer surfactant corona dynamically replaces water in solvent-free protein liquids and ensures macromolecular flexibility and activity. *J. Am. Chem. Soc.* **134**, 13168–71 (2012).
 25. Costantino, H. R., Curley, J. G. & Hsu, C. C. Determining the water sorption monolayer of lyophilized pharmaceutical proteins. *J. Pharm. Sci.* **86**, 1390–1393 (1997).
 26. Tonlolo, C. & Benedetti, E. The polypeptide 310-helix. *Trends Biochem. Sci.* **16**, 350–353 (1991).
 27. Barlow, D. J. & Thornton, J. M. Helix geometry in proteins. *J. Mol. Biol.* **201**, 601–619 (1988).
 28. Brogan, A. P. S., Siligardi, G., Hussain, R., Perriman, A. W. & Mann, S. Hyper-thermal stability and unprecedented re-folding of solvent-free liquid myoglobin. *Chem. Sci.* **3**, 1839 (2012).
 29. Moriyama, Y. *et al.* Secondary structural change of bovine serum albumin in thermal denaturation up to 130 degrees C and protective effect of sodium dodecyl sulfate on the change. *J. Phys. Chem. B* **112**, 16585–16589 (2008).

Chapter 4 Protease activity in the solvent-

free liquid state

4.1 Introduction

Protease activity, or proteolysis, is the action of a protease catalysing the hydrolysis of a peptide bond. Proteolysis is known to be crucial in a multitude of physiological regulatory processes within cells, *e.g.* apoptosis¹. Therefore, any alteration to the natural activity can have a potentially detrimental impact on the host organism.

Proteases are a family of enzymes (detailed in Section 1.1.6) that are increasingly being used to catalyse industrial processes. The wide range of available proteases means there are a large number of different reactions where it can be usefully employed. However, the high temperatures and pressures of many syntheses is a current limiting factor for which proteases can be used due to the resulting denaturation and activity retardation. As such, protease activity needs to be closely monitored and assessed using enzyme activity assays. Preferably, these assays would be universally designed and therefore compatible to work with a large range of proteases and their substrates so that activity comparisons can be made.

4.1.1 Protease activity assays

To follow the activity of a protease, either the cleaving of the substrate must directly alter a detectable signal, or a detectable signal is generated through the interaction of the cleaved peptide chains with non-protease/non-substrate molecules. Protease activity assays are often classed as either homogeneous or heterogeneous. A homogeneous assay is one where the protease and substrate, along with any additional tracking molecules are in the solvent-phase. Homogeneous assays often take the form of calorimetric, mass spectrometry, UV-Vis spectrometry, or fluorescent resonance energy transfer (FRET) based assays². Conversely, a heterogeneous activity assay has one of either the protease or substrate within the solvent phase and the other is immobilized on an inert substrate. Often, the species which is immobilized is dependent on which is most stable to denaturation within the required solvent; proteases generally requiring solvation within water for conformational stability. Heterogeneous assays often take the form of fluorometric, electrochemical, or surface plasmon resonance-based assays³.

Another protease assay classification criterion is whether the activity can be tracked continuously throughout the reaction period. A continuous protease activity assay involves the cleaving of the peptide bond directly generates a detectable signal, or where the interaction of a cleaved peptide fragment with an additional molecule within the system generates a detectable signal. Often a continuous assay would be selected when the specific rate of proteolysis is required or if there are designated times when the reaction is perturbed either through an environmental or chemical fashion that could alter the rate. An example of a continuous assay would be where a known FRET (fluorescence resonance energy transfer) pair of fluorophores is placed either side of a peptide bond.

Cleavage of the bond separating the FRET pair directly alters the fluorescent signal⁴. Conversely, a discontinuous activity assay involves the cleaving of the peptide bond or the interaction of the peptide fragment with an additional molecule does not directly generate a detectable signal, rather the solution is analysed at the end of the reaction period. Discontinuous assay would be selected when the rate of proteolysis is not required for the system, when only maximum cleavage is required, or more simply the detection of cleavage is needed. An example of a discontinuous assay would be one where an additional molecule is added post-reaction, then the signal generated from the interaction of this molecule with the peptide fragments in the sample can be analysed. It is interesting to note that if the reaction can be quenched, discontinuous assays can be pseudo-continuous, in that a low temporal resolution can be achieved. The selection of protease activity assay is therefore dependent on three factors: the solubility of the components within the assay, the required detection method, and also desired outcome.

4.1.1.1 Enzyme assay requirements

The choice of enzyme activity assay for this study was based on the need to assess how the various stages of the solvent-free liquid protein synthesis upon a protease (proteinase K (PK)) and a protease substrate (*bovine serum albumin* (BSA)) affected the proteolytic activity. Each stage of the synthesis for each protein was assessed separately to discover how the surface modification altered the protease activity, with subsequent assessment of how pairings of modifications altered the protease activity. The final activity assay of the set is the pairing of the solvent-free liquid PK ($[C-PK][S_x]$) and solvent-free liquid substrate (where x is either 1 or 7 (Chapter 2). Figure 4.1 is a graphical representation of the required activity assays.

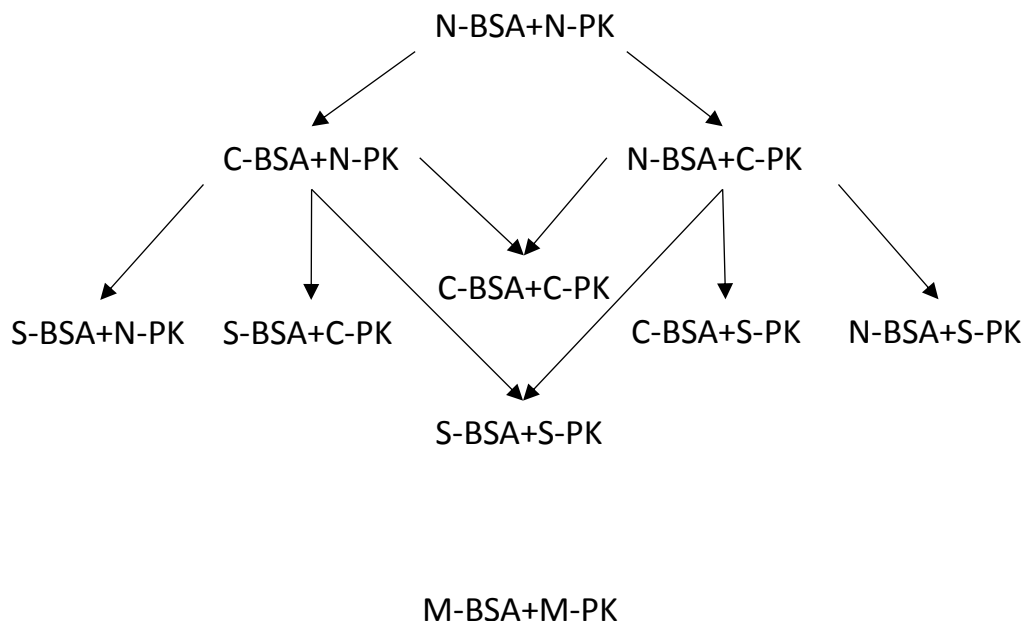


Figure 4.1: A graphical representation of the required activity assay pairings where; N is the native species, C is the cationized species, S is the surfactant conjugated species, and M is the solvent-free liquid protein species, and PK is proteinase K, and BSA is *bovine serum albumin*.

The two requirements for the protease activity assay for the protease and protease substrate pairings that were investigated in the following chapter were that the assay must be homogeneous and continuous. This allowed for PK and substrate to be solvated in the aqueous phase and for the activity to be continuously assessed in order to obtain proteolytic activity rates, thus allowing for comparison between systems. The assay was also based around the protease substrate, the protein that was subjected to the solvent-free liquid protein synthesis and allowed for the solvent-free liquid phase of the protease and substrate to be assayed along with the precursor molecules. An additional requirement for the assay was to be of a high throughput, allowing for numerous runs to be conducted in a short time period due to the large number of assay pairs that were investigated in this study. The following is a short discussion on the development and selection of the assay, detailing two potential but unused assays and the reason why they did not meet the requirements.

4.1.2 Common protease activity assays

A common assay used for the detection of protease activity is a non-specific protease casein-substrate based assay⁵. This assay can be used for a wide range of proteases as the substrate does not have a tightly folded conformation, leaving a large number and range of different potentially cleavable peptide bonds exposed, and is also amphiphilic, displaying many hydrophilic and hydrophobic residues. The assay functions through the selected protease cleaving the casein throughout the reaction period with tyrosine (in addition to many other residues) being released either as individual

residues or as peptide fragments, this being protease specific. Post-reaction, a Folin and Ciocalteus phenol reagent reacts with free tyrosine and produces a blue chromophore and is therefore detectable *via* UV-Vis spectroscopy. It then follows that the higher the activity of the protease the more tyrosine is liberated from the casein, and consequently more chromophores are produced giving a higher UV-Vis absorbance signal. This assay would be classified as homogeneous and discontinuous as the protease and substrate are both within the aqueous phase and the assessment of activity is conducted once at the end of the reaction. As the assay is discontinuous it did not meet the predefined requirements for the PK-BSA assay.

Fluorimetry is another common technique for tracking protease activity, with the protease either cleaving peptide bonds that unquench/generate a fluorophore with, activity being indicated *via* an increase in fluorescence intensity, or where the cleavage removes the fluorescent nature of a fluorophore, thus the activity is indicated by a decrease in fluorescence intensity. An ideal activity assay for this system would be for the protease substrate to be an inherently fluorescent protein, *e.g.* green fluorescent protein (GFP) or mCherry. In this case, the benefits would be two-fold; firstly, the surface modifications that occur during the solvent-free liquid protein synthesis would alter the fluorescence of the protein and this could then be used as a method for assessing how the protein has changed due to the synthesis; secondly, the fluorescence could be used as an indicator of the protease activity, with a decrease in fluorescence being directly linked to the protease activity. However, it is well documented that fluorescent proteins such as GFP are resistant to proteolysis by proteases, despite numerous putative sites being displayed on highly exposed surface loops⁶. Investigations into synthesising a more protease susceptible GFP have been successfully conducted through the insertion of 5 amino acid residues to a number of these highly exposed surface loops. The GFP mutants in this study were shown to be susceptible to a number of proteases exhibiting the possibility for fluorescence tracking. However, these assays were conducted over 12 hour time periods, with only a 40% drop in fluorescence in that time. This assay would be classified as homogeneous and continuous, yet due to the long time period it does not meet the predefined requirements for the PK-BSA assay.

Numerous other assays were considered for the study of how the solvent-free liquid protein synthesis affected the activity of a protease towards its substrate, however, they did not meet all the predetermined requirements. Therefore, it was decided to create and develop a novel assay to track the activity of the protease. The key benefit with the developed assay being that it was designed to be applicable to both the aqueous precursor constructs as well as the solvent-free liquid protein phase. In brief, the assay utilises circular dichroism to track the protease degradation of the substrate's secondary structure, discussed in more detail in the following section.

4.2 The development of a novel circular dichroism spectroscopy-based protease assay

4.2.1 CD assay introduction

All proteins have a unique secondary structure that is retained through a delicate balance of forces and non-covalent bonds that oppose the entropically favoured unfolding of a protein. This secondary structure can be observed through circular dichroism (CD) spectroscopy as detailed in Chapter 2.1.1. Often unmentioned when discussing the stability of a protein's secondary structure are the covalent peptide bonds that hold the proteins primary sequence together. Within a protein if a peptide bond is broken, the balance of forces stabilising the protein is drastically shifted and can immediately induce a large secondary structure rearrangement. This alteration in secondary structure can be observed and tracked instantaneously and continuously through CD spectroscopy.

Through monitoring the structural changes detailed above, a novel circular dichroism-based assay was developed to continuously track the activity of a protease, proteinase K (PK), and a protease substrate, *bovine serum albumin* (BSA), both in the aqueous phase and in the solvent-free liquid protein phase. CD spectroscopic study of a protein is generally conducted in the aqueous phase, however, if a sample is a clear viscous liquid then solid-state CD spectroscopy can be conducted on the sample. Enzyme kinetic models can be used to describe and compare the activity of multiple systems because the tracking of the secondary structure change is continuous. This section details a novel CD-based protease activity assay in general terms for any protease substrate pairing, rather than being specific for PK and BSA, and how an enzyme kinetic model was fitted, along with controls deemed necessary to prove the validity of the assay.

4.2.2 Assay functionality

Circular dichroism spectroscopy (as detailed in Sections 2.1.1) is an optical spectroscopic technique that is used to detect secondary structure in proteins. Any perturbation to the conformation, either through environmental change or a destabilization event, can disrupt the secondary structure motifs altering their optical activity, which in turn alters the obtained spectrum of the protein.

In the context of this protease assay, the initial destabilization event is the proteolytic cleavage of a peptide bond displayed on the surface of the substrate protein. This initial cleavage has the potential to occur in many places on the surface, however, each protease will have preferential initial cleavage sites on each substrate protein based on the specificity of the individual protease. Depending on where the initial cleavage occurred determines the extent of the initial secondary structure alteration and subsequently the CD signal, *i.e.* if the protease merely cuts a surface loop the impact may be small relative to the cleavage being within an α -helix, which is likely to disrupt the whole helix and drastically

change the optical properties of the protein. The initial cleavage, and the following structural change, also has the potential to reveal areas of the structure that would not have been initially accessible to the protease. This means that subsequent cleavage events are not limited to the surface accessible peptide bonds. Each cleavage event that occurs will disrupt the secondary structure, further degrading the CD signal of the substrate protein up until the point when that specific protease has no more cleavage sites available. At this point, the secondary structure of the substrate protein has been degraded as far as the protease can, observed as a plateau in the CD signal. Each individual cleavage event is not distinguishable since the CD signal detected is an average over the sample population. However, each protease will have a preferential cleavage pathway for the substrate protein, meaning that each substrate protein, on average, will be cleaved in the same fashion.

4.2.3 Assay development: Controls and preliminary results

The assay is designed to follow the proteolytic degradation of a substrate proteins secondary structure over time through the detectable decline in CD signal throughout the reaction. By plotting the decline in CD signal of one or more of the characteristic peaks, *e.g.* 222 nm for α -helical dominant proteins, the rate of the protease activity can be calculated. For the assay to be tracking the cleavage of the substrate protein, the substrate must be in excess in relation to the protease, ensuring that the observed CD signal is solely related to secondary structure of the substrate and not that of the protease. It should be noted that the decline in rate of cleavage observed for a sample is not solely a result of the decline in available substrate cleavage sites. As the protease is also a protein there is the potential that it will also be cleaved by other proteases in the system. Any cleavage of the protease will retard the rate due to conformation changes to the active site, potentially rendering the protease in-active. Over time, the number of active proteases will potentially decline, with the rate of the observed CD signal degradation declining as a result. However, as the rate of protease 'self-degradation' is consistent over all assays, the degradation rates between substrate proteins are still comparable, with the change in observed rates between samples being exclusively related to the different substrate concentrations. It is also clear from Figure 4.2 A and B that proteinase K is resistant against self-catalysis (there being no change in spectrum over 60 minutes), meaning the rate of proteolysis won't slow down due to a loss of proteases.

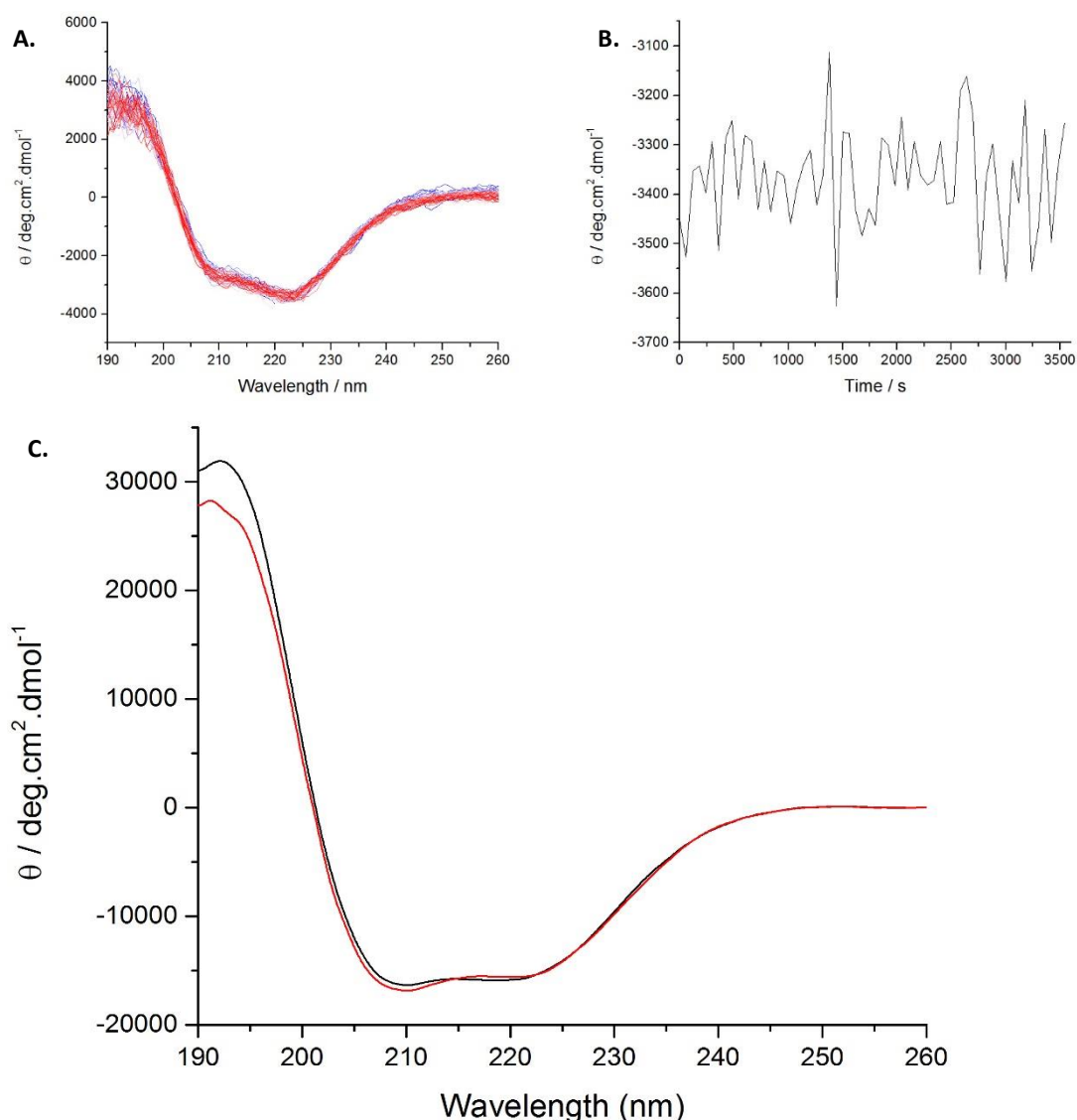


Figure 4.2: CD spectra of **A.** PK v PK over 60 minutes (red to blue indicating reaction time) [concentration of PK is 0.2 mg/mL⁻¹] **B.** 222 nm CD spectrum of PK v PK over 60 minutes [concentration of PK is 0.2 mg/mL⁻¹], and **C.** CD spectra of native-BSA (black) and BSA and PK (red), where the concentration of BSA is 0.2 mg mL⁻¹ for both spectra and the concentration of PK is 0.02 mg mL⁻¹.

It is interesting to note that the CD spectrum observed for the reaction between the protease and protease-substrate may not be equivalent to the spectrum observed for the individual substrate (Figure 4.2). When a protease interacts with the substrate, the secondary structure of the substrate is likely to alter, the binding inducing an observable conformational change and the formation of a protease-substrate complex. As the concentration of the protease is too low to be directly detected, this alteration to the average CD signal relative to the individual substrate is caused by the formation of protease-substrate complexes, with the concentration of complexes being equal to or lower than that of the concentration of the protease. However, it is clear from comparing the CD spectra of native-BSA and native-BSA with native-PK that they are very similar.

Taking these factors into consideration, the concentration of the protease must therefore be low enough to not be detectable with CD spectroscopy, low in relation to the substrate to ensure minimum collisions between other proteases within the system, but also at a concentration that is high enough to generate a rapid and large signal decrease, *i.e.* enough substrate cleavages per second to be detectable as a CD signal change. The time period of each reaction pairing is dependent on the speed of degradation, for pairings that have a relatively low activity the time taken to acquire the CD spectra will be longer. Below is an example of one protease/protease-substrate pairing, using PK and BSA as the example protease and protease-substrate, respectively.

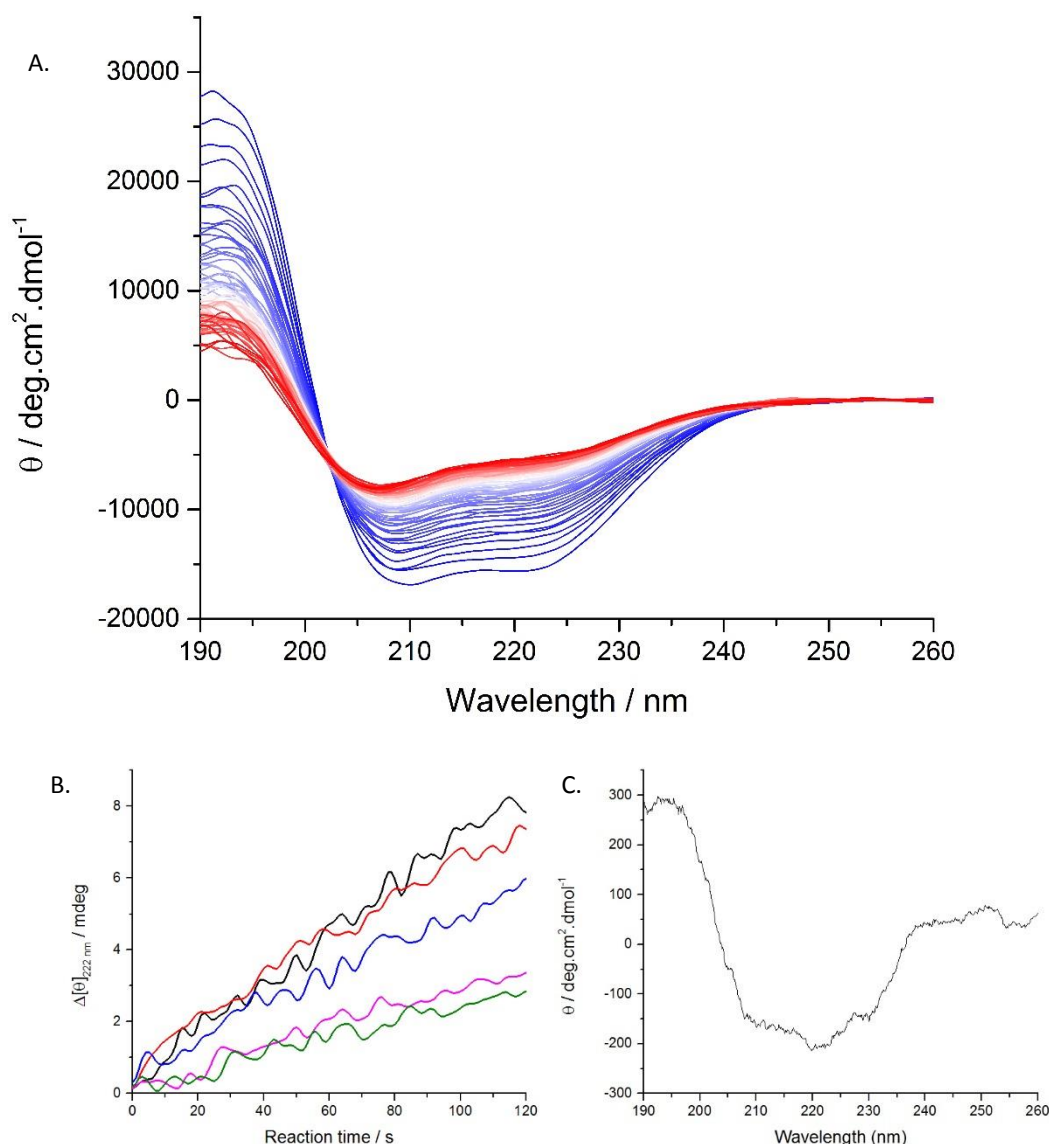


Figure 4.3: **A.** CD spectra of BSA and PK over a period of 30 minutes (reaction time moving blue to red) where the concentration of BSA and PK are 0.2 mg.mL⁻¹ and 0.02 mg.mL⁻¹, **B.** 222 nm CD wavelength of 3 BSA and PK spectra taken over a period of 2 minutes [Concentration of BSA is 0.0129 mM (black), 0.0088 mM (red), 0.0068 mM (blue), 0.0027 (magenta), and 0.002 (green)], and **C.** the CD spectrum of native-PK at a concentration of 0.02 mg.mL⁻¹.

It is evident that as the reaction proceeds there is a rapid decrease in CD signal, with this decrease being replicated in repeat assays (Figure 4.3A). BSA is known to have a high percentage of α -helices, therefore, by plotting the change in 222 nm wavelength (a wavelength strongly associated with α -helical structures) against time, the initial rate for this system can be calculated. This initial rate of activity can then be used to compare the rates of multiple systems (Figure 4.3B). The concentration of PK is kept low to reduce the impact that it has on the CD spectrum (Figure 4.3C).

4.2.3.1 *Initial spectra shape of pairings relative to lone substrate*

It was important to assess whether the concentration of the protease was low enough to not be directly or indirectly (through the formation of complexes) detectable in each of the pairings. This was done by the spectra of the pairings at reaction time = 0 being compared to the spectra of the relevant substrate, and the concentration of PK used, $0.69 \mu\text{M}^{-1}$, being analysed *via* CD. If the spectrum for PK at the lowest selected concentration had a very high signal to noise ratio then this would imply that the concentration is too low to be effectively detected via CD. If the spectra for the initial stages of the reactions were similar to the individual substrate spectra, this would indicate that the ratio of BSA:PK was largely in favour of BSA, showing that the spectrum was not affected by any PK/BSA complexes formed.

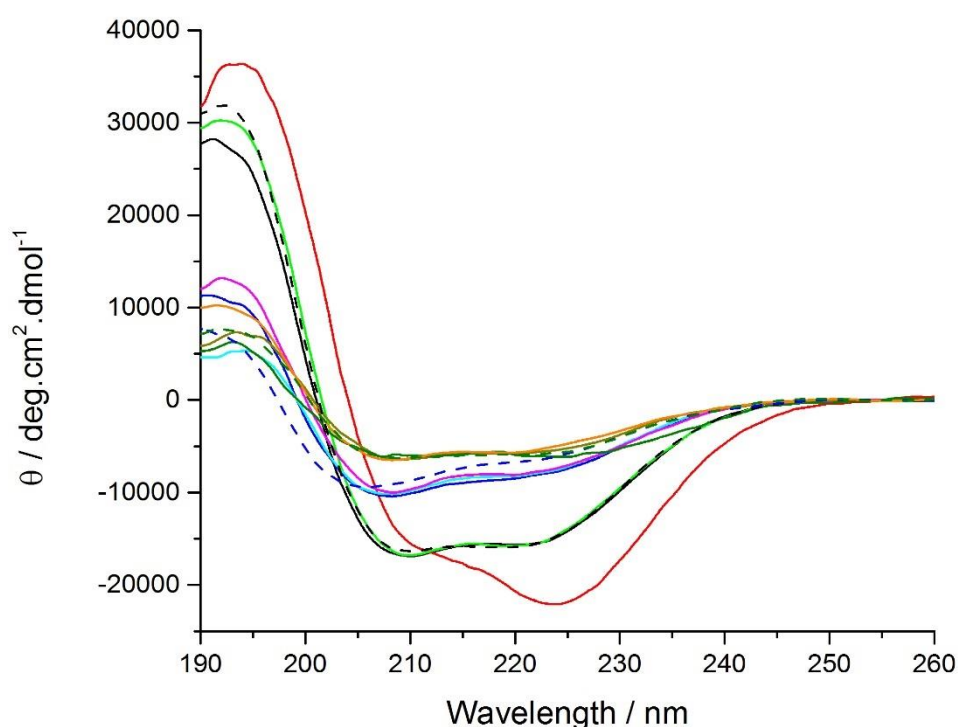


Figure 4.4: Mean residue ellipticity against wavelength showing far-UV CD spectra for NBSA (dashed black), C-BSA (dashed blue), and aqueous [C-BSA][S₇] (dashed dark green), along with spectra at time = 0 s for the following reactions: NBSA vs NPK (black), NBSA vs C-PK (red), NBSA vs aqueous [C-PK][S₇] (light green), C-BSA vs NPK (blue), C-BSA vs C-PK (cyan), C-BSA vs aqueous [C-PK][S₇] (magenta), aqueous [C-BSA][S₇] vs NPK (dark yellow), aqueous [C-BSA][S₇] vs C-PK (dark green), and aqueous [C-BSA][S₇] vs aqueous [C-PK][S₇] (orange).

It is clear from the spectrum of PK that no usable reproducible spectrum is obtainable at this low concentration (Figure 4.3C). All the protease/substrate pairings (Figure 4.4) at a PK to BSA number ratio of 1:4, with the exception of NBSA v C-PK (red), displayed retention of the individual relative BSA secondary structure, be that BSA, C-BSA, or aqueous [C-BSA][S₇]. This indicates that the concentration of PK relative to BSA, even at this low ratio, was low enough not to alter the average observed structure to a large degree through the formation of PK/BSA complexes.

The initial spectra for NBSA v C-PK (red) does not match the spectrum of NBSA, displaying red-shifted 208 and 222 nm characteristic wavelengths, to 210 and 224 nm respectively (Figure 4.4). There is also a large change in the $\theta_{222}/\theta_{208}$, from being ≈ 1 for NBSA to 1.4 for NBSA v CPK. This type of red-shifting in the CD spectrum has been known to occur due to absorption flattening, an effect known to manifest from samples that are aggregated or are inhomogeneous⁷. The increase in the $\theta_{222}/\theta_{208}$ ratio has been attributed to the interlacing of helices. In this context, this could be manifested through the forming of quaternary protein complexes or tightly bound aggregates, potentially comprised of either PK, BSA, or PK+BSA. From this evidence it would seem apparent that the sample, to at least some extent, was initially aggregated or partially complexed. As this effect only occurred between this pairing, a more thorough investigation would need to be conducted to ascertain the cause of aggregation. The

formation of C-PK, as previously detailed, does alter the secondary structure, and although it does not aggregate when interacting with other C-PK proteins, or even with C-BSA and aqueous [C-BSA][S₇], the interaction with NBSA does cause some form of aggregation leading to interlaced helices.

Enzyme kinetic models are used to describe the kinetic properties of an enzyme reaction, allowing for more a well-defined qualitative analysis of an enzyme system. The Michaelis-Menten kinetic model was selected and is detailed in section 2.1.3. The model yields values for the maximum enzymatic rate of a system and the enzyme binding affinity, the basis for the interpretation of how the surface modifications alter the activity. This model requires the concentration of the protease to be kept constant whilst the concentration of the substrate is increased.

4.3 Results for the aqueous precursors

Each aqueous pairing of PK and BSA was investigated through the use of the previously detailed protease activity assay, with the Michaelis-Menten enzyme kinetic model being used to calculate the maximum rate of reaction and the Michaelis-Menten constant (an analogue for the substrate binding affinity). For each pairing, a representative full CD wave scan of the reaction along with graphs of the initial rate against concentration are displayed in the following section.

4.3.1 CD assay discussion

As previously discussed in section 4.2.3.1, the spectrum for the majority of the reactions should initially be dominated by a BSA-like spectrum as the protease concentration is at an undetectable level. This also means that the complex formed between the BSA and the PK should be relatively undetectable in most cases, since the number of potential complexes here was limited by the number of proteases present. For the systems herein, the concentration of PK was set at $0.69 \mu\text{M}^{-1}$, with the concentration of BSA ranging from 0.75 - $33.1 \mu\text{M}^{-1}$.

Each protease in the pairing was subject to a number of different substrate concentrations, with each concentration repeated in triplicate. Preliminary experiments demonstrated that when the surfactant variant of either the protease or the substrate was a member of the pairing the experimental duration, for the same amount of structural degradation to be observed, could take at least ten times longer. Because of this, it was decided that these pairings would only be assessed at 5 different substrate concentrations, with all other pairings being assessed at 20 different substrate concentrations. All calculated numerical values are in Table 4.1.

4.3.1.1 Degradation pathway

It is interesting to note that the majority of the pairings have relatively similar degradation curves at all three of the characteristic peaks (195 nm, 208 nm, and 222 nm) (example in Figure 4.6). This could indicate that degradation of the substrate's secondary structure moieties occurs relatively consistently throughout the sample, with the rate of change of wavelength at similar times throughout the reaction (Figure 4.6). This is true for 8 out of the 9 pairings, with aqueous [C-BSA][S₇] v C-PK being the outlier.

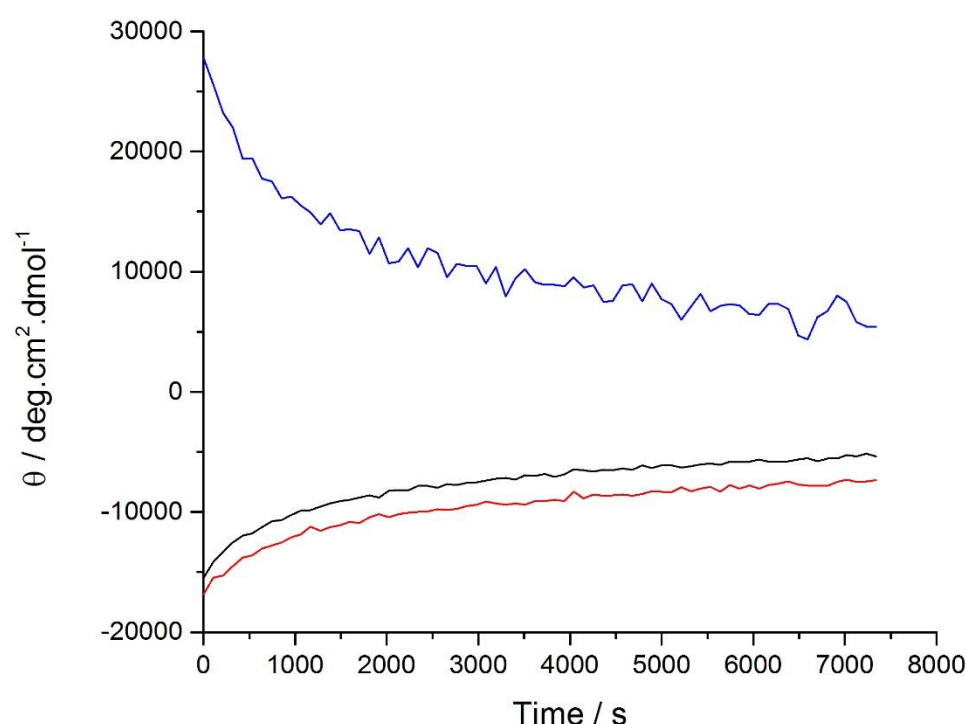


Figure 4.5: Mean residue ellipticity against wavelength showing far-UV CD spectra for NBSA v NPK at wavelengths of 195 nm (blue), 208 nm (red), and 222 nm (black).

Although the initial spectra of aqueous [C-BSA][S₇] v C-PK is analogous to aqueous [C-BSA][S₇] it is clear from the full spectrum scan of the reaction (Figure 4.6) that there is a rapid and drastic alteration to the observed secondary structure. There is an abrupt conformation shift after 20 minutes from being an α -helical dominant spectrum, displaying characteristic negative 208 and 222 nm peaks along with a positive 195 nm peak, to being fully random coil in nature, displaying a characteristic large negative 195 nm peak and a positive 222 nm peak. This indicates that after 20 minutes, on average, the C-PK begins to cleave structurally crucial peptide bonds within aqueous [C-BSA][S₇], initiating the structural denaturation of the protein. After this event, the degradation of the structure is observed at a relatively lower rate due to the larger, more optically active motifs now being absent from the sample. As no other pairing displayed this drastic shift, it implies that it is not solely related to the modification

of either individual protein, and more a result of the modification combination. The specific mobile molten globular nature and surfactant hinderance of aqueous [C-BSA][S₇] paired with the likely altered conformation of the active site of C-PK could change the degradation pathway of this pairing in relation to aqueous [C-BSA][S₇] v PK. Due to the initial cleavages differing from the other pairings it is possible that these cleavages exposed internalised hydrophobic residues that then interacted with the hydrophobic domains in the surfactant, initiating more unfolding of the protein, a drastic effect not seen in other surfactant conjugate pairings. This combination of surface alterations leads to the observed radical degradation of the substrates structure, with it being the only pairing to have an effect to this extent. The kinetic values for this pairing were incalculable using the Michaelis-Menten kinetic model as no curve could be fitted to the data.

4.3.1.2 V_{max}

Table 4.1: The maximum rate of activity (V_{max}) and binding affinity (K_M) for all pairings of PK, C-PK, aqueous [C-PK][S], BSA, C-BSA, and aqueous [C-BSA][S].

V_{max} ($\Delta mdeg.s^{-1}$)	NPK	C-PK	Aqueous [C-PK][S ₇]
K_M (mM)			
NBSA	128±26	35.4±2.72	1.58±0.98
	0.059±0.015	0.004±0.001	0.011±0.005
C-BSA	104±23	4.48±2.36	0.94±0.13
	0.49±0.12	0.051±0.039	0.003±0.001
Aqueous [C-BSA][S ₇]	5.86±2.42	NA	0.67±0.13
	0.024±0.014	NA	0.003±0.001

4.3.1.2.1 Cationization of PK

Cationization of PK was previously shown to alter the secondary structure, increasing the α -helical percentage by 8% (Section 3.2.2.3.3). This shift in structural conformation has a high probability of affecting the structure of the active site. Any changes to the configuration of the active site can alter the relative position of the binding cleft to the catalytic triad, or even the orientation of the catalytic triad, both of which could retard or eliminate the protease activity. Cationization also has the potential to place additional positive charges proximal to binding cleft residues modifying how the protease electrostatically interacts with the substrate, conceivably disrupting the catalytic triad or introducing repulsion to the overall interaction potential. As there are no solvent-accessible sites that can be cationized within the active site of PK⁸, this cannot be a cause for the protease to lose activity. Nevertheless, the global increase in surface charge on the protease will introduce additional electrostatic repulsion at the surface of BSA, increasing the energy required for the two proteins to catalytically interact, thus reducing the activity. This additional repulsion, along with the known

structural change, makes it highly likely that the cationization of the protease will reduce the activity towards each BSA substrate relative to NPK.

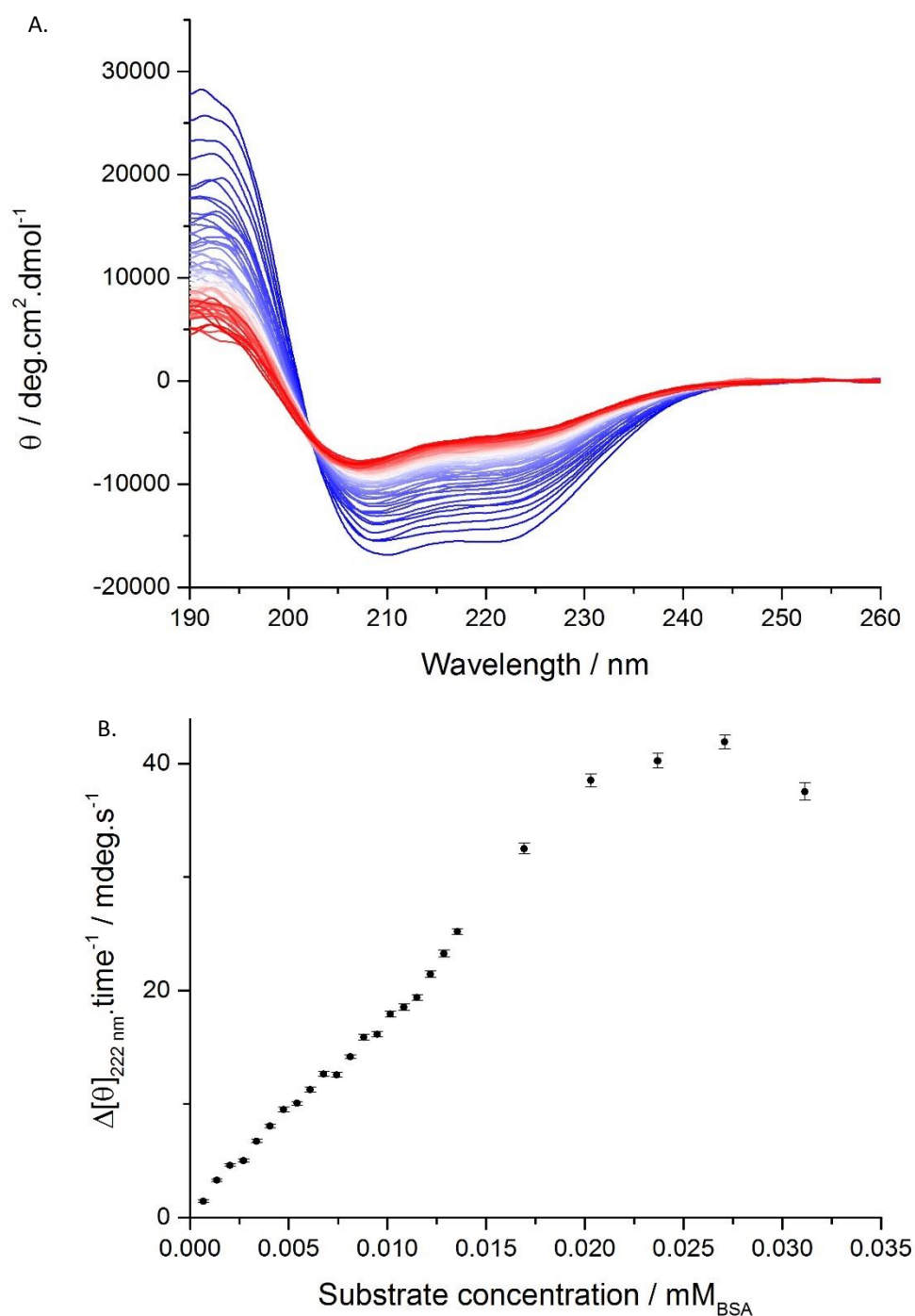


Figure 4.6: **A.** Mean residue ellipticity against wavelength showing far-UV CD spectra for the NBSA vs NPK reaction in the range of 190–260 nm over a time period of 120 minutes with a full scan every 2 minutes (reaction time progressing from blue to red) [concentration of BSA 0.0027 mM], **B.** Initial rates of reaction against substrate concentration for the NBSA vs NPK reaction (each data point representing the average of a minimum of 3 repeats).

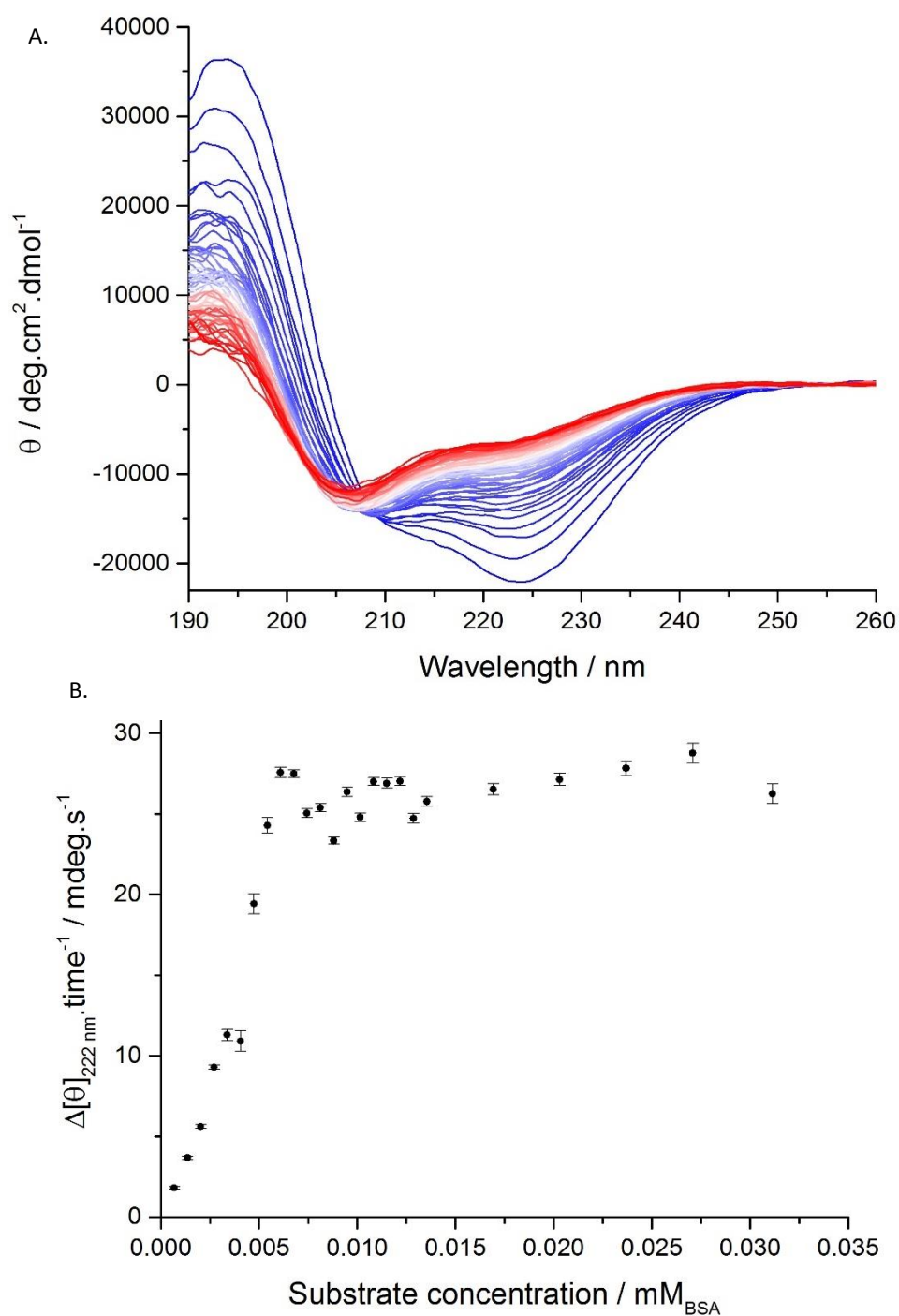


Figure 4.7: **A.** Mean residue ellipticity against wavelength showing far-UV CD spectra for the NBSA vs C-PK reaction in the range of 190 – 260 nm over a time period of 120 minutes with a full scan every 2 minutes (reaction time progressing from blue to red) [concentration of BSA 0.0027 mM], **B.** Initial rates of reaction against substrate concentration for the NBSA vs C-PK reaction (each data point representing the average of a minimum of 3 repeats).

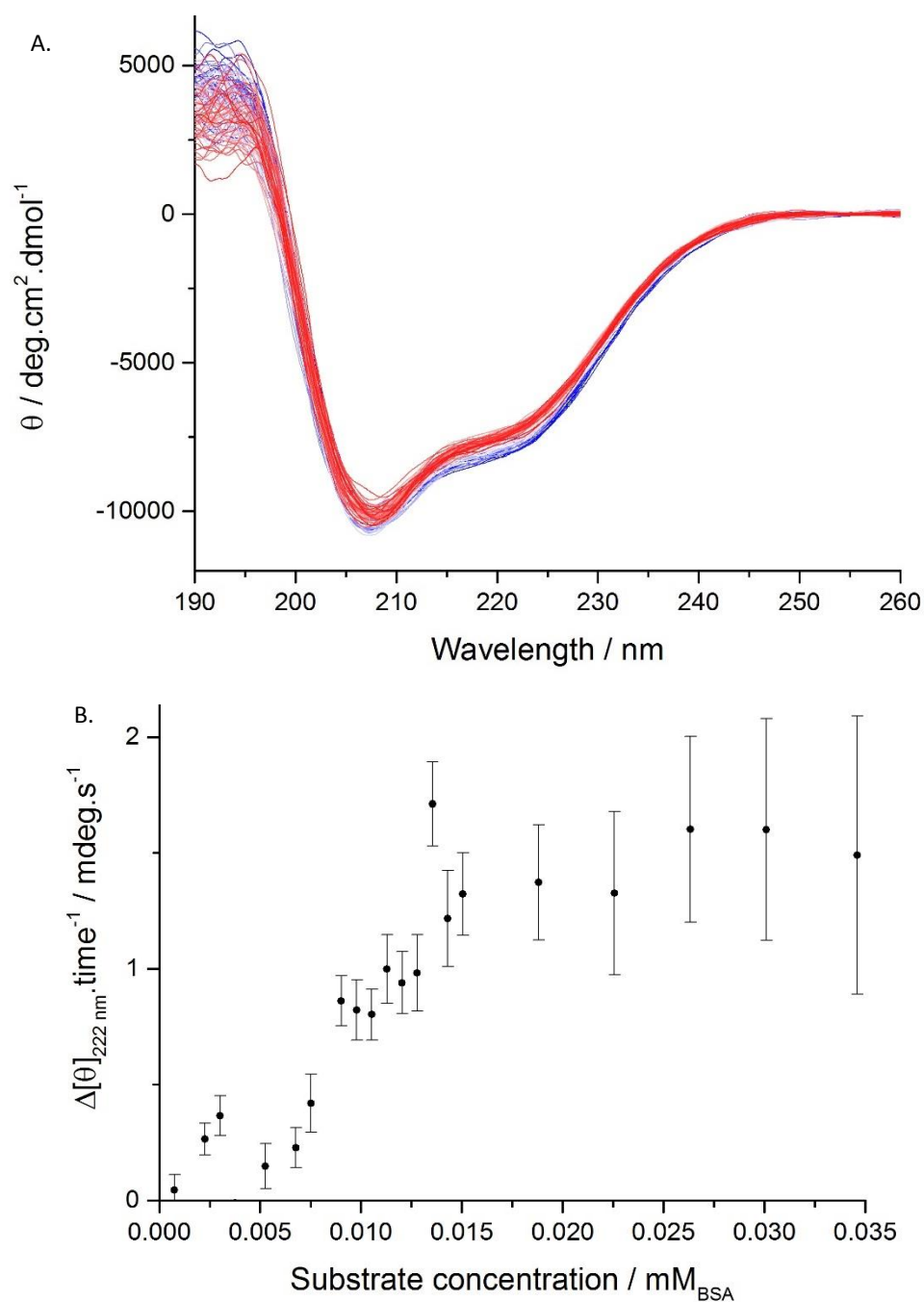


Figure 4.8: **A.** Mean residue ellipticity against wavelength showing far-UV CD spectra for the C-BSA vs C-PK reaction in the range of 190 – 260 nm over a time period of 2300 minutes with a full scan every 4 minutes (reaction time progressing from blue to red) [concentration of BSA 0.0027 mM], **B.** Initial rates of reaction against substrate concentration for the C-BSA vs C-PK reaction (each data point representing the average of a minimum of 3 repeats).

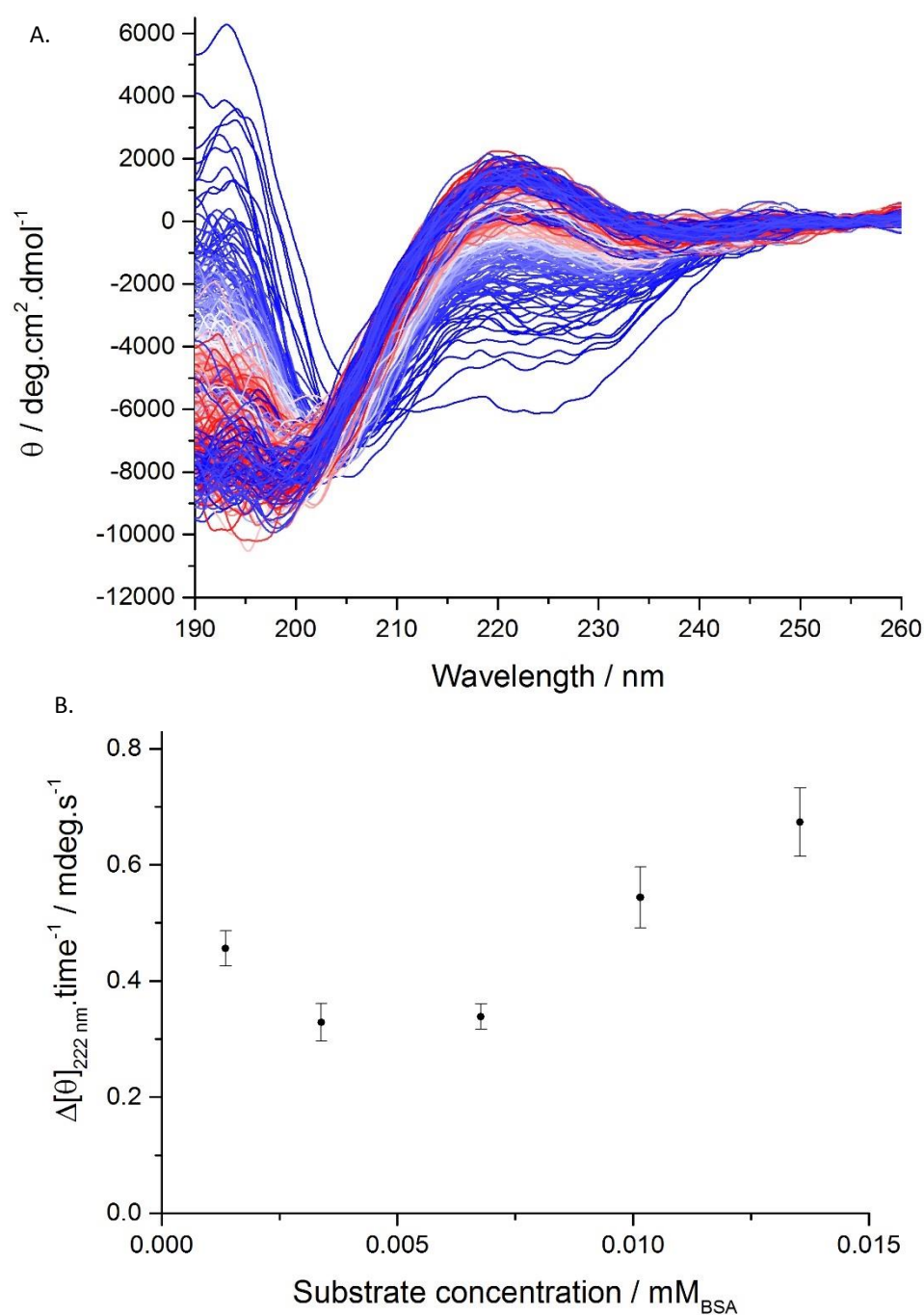


Figure 4.9: **A.** Mean residue ellipticity against wavelength showing far-UV CD spectra for the aqueous [C-BSA][S₇] vs C-PK reaction in the range of 190 – 260 nm over a time period of 1000 minutes with a full scan every 4 minutes (reaction time progressing from blue to red) [concentration of BSA 0.0027 mM], **B.** Initial rates of reaction against substrate concentration for the aqueous [C-BSA][S₇] vs C-PK reaction (each data point representing the average of a minimum of 3 repeats).

In comparison to NBSA, the cationization of PK to form C-PK had a large negative impact on the activity of the protease towards the substrate, with a 73% drop in V_{\max} relative to NBSA v NPK (Figure 4.6). As previously discussed, the retardation in activity rate exhibited here is likely due to the conformational alteration of the active site as a consequence of the surface modification, along with an increase in overall protease/substrate electrostatic repulsion. Since the altered binding cleft displays interaction sites in different locations, the binding now takes increased energy to complete. This energy required is further increased due to also overcoming the repulsive forces.

In relation to C-BSA, the cationization of PK to form C-PK had a detrimental impact on the activity of the protease towards the substrate, with a 96% drop in V_{\max} relative to C-BSA v NPK (Figure 4.7). The large loss of activity displayed in this pairing is attributed to a combination of loss of protease structure leading to a loss in activity and the increase in electrostatic repulsion, also paired with the idea that C-BSA aggregates take a longer time period to degrade. As both PK and BSA are cationized, with a net increase of 30 and 200 positive surface charges respectively, there is a large increase in electrostatic repulsion between the two proteins. This electrostatic repulsion will decrease the probability of protein-protein interaction and therefore greatly hinder the activity.

Due to the data from the kinetic assays run on aqueous [C-BSA][S₇] nanoconjugate v C-PK not fitting the Michaelis-Menten kinetic model (Figure 4.9), the V_{\max} for this pairing is unattainable (discussed in Section 4.3.2.1). Although, as there was degradation of the surfactant conjugated substrate it displays that surfactant conjugation does not prohibit proteolytic action. While the method of proteolytic action on a surfactant covered protein surface remains unknown, a mechanism is postulated here. It is possible that some areas of the surface that are not coupled to surfactant molecules are the long-extended regions of the substrate that the protease can initially interact with, since proteases are known to not interact with folded secondary structure elements of a protein⁹. In addition the more unfolded molten-globular [C-BSA][S₇] nanoconjugate would display more of these extended unstructured regions. Therefore, the coupling of these two effects; lack of surfactant coverage on protease active areas and the molten globular state affording more mobility of these areas, could explain the protease activity towards a surfactant conjugated protein.

The successful, albeit retarded, degradation of N-BSA, C-BSA, and aqueous [C-BSA][S₇] using C-PK demonstrated that despite the likely alteration to active site conformation and the large increase in positive surface charge, protease activity was retained.

4.3.1.2.2 Surfactant conjugation of PK

Surfactant conjugation of PK was previously shown to reclaim some of the native secondary structure of PK (Section 3.2.2.3.3), decreasing the α -helical content by 8% and the β -sheet content by 7% relative to C-PK. As with cationization, any changes to the configuration of the active site brought upon by the shift in conformation can alter the interaction of the protease with the substrate, subsequently retarding or inhibiting the protease activity. The addition of anionic surfactant molecules to positively charged surface sites of C-PK is thought to partially shield the many positive charges on the surface not only from each other, thus reducing surface electrostatic repulsion, but also from repulsive interactions with neighbouring proteins. The polymer chains are known to cover the surface of the protein¹⁰, with the tail sections potentially entering the active site of the protease through random walk movement. This has a high probability of disrupting interactions between the active site and a substrate, with the possibility of creating steric hindrance at the active site and preventing protease-substrate interaction. This steric hindrance, along with the known structural change, makes it highly likely that the surfactant conjugation of the protease will reduce the activity towards each BSA substrate, relative to NPK and C-PK.

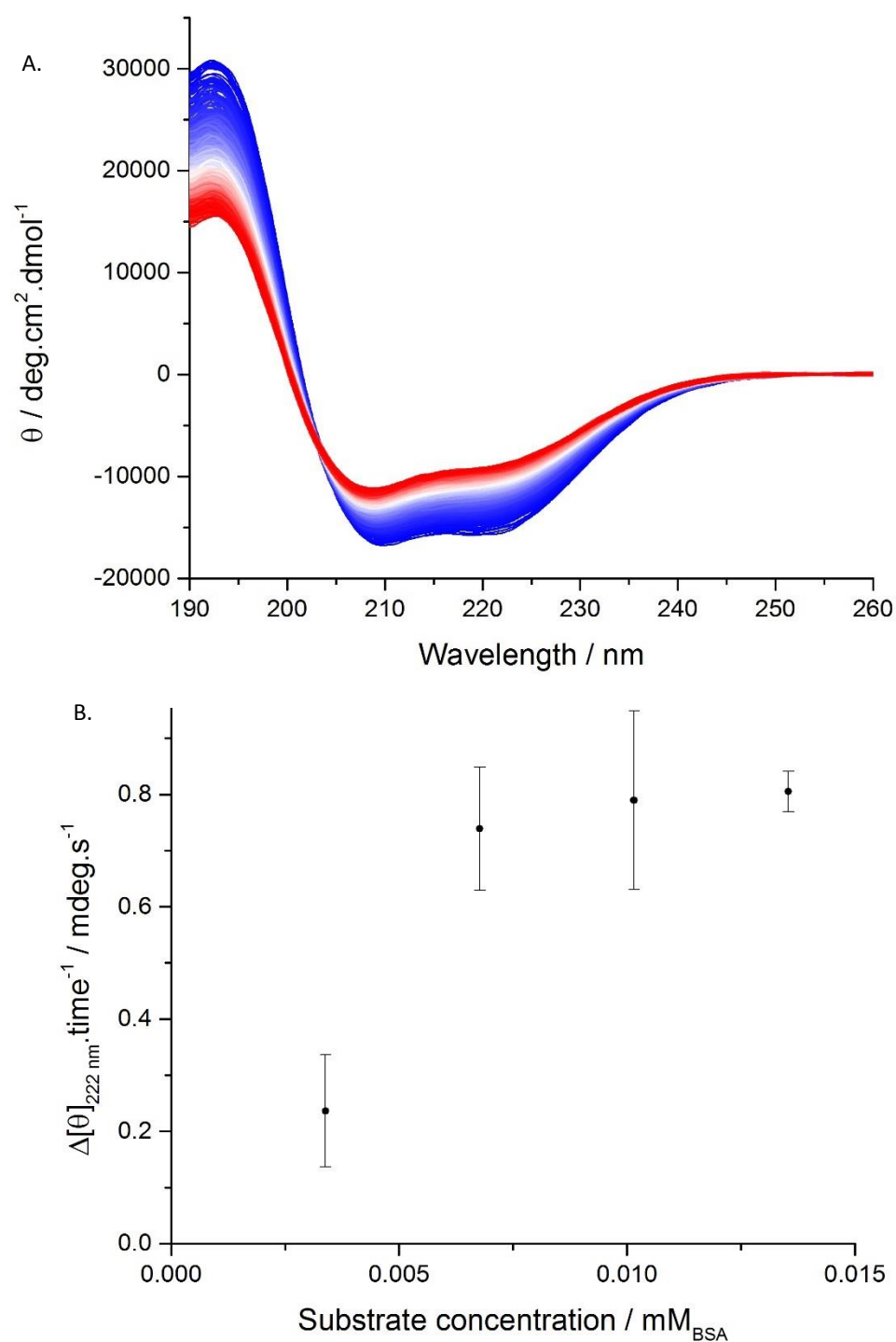


Figure 4.10: **A.** Mean residue ellipticity against wavelength showing far-UV CD spectra for the NBSA vs aqueous [C-PK][S₇] reaction in the range of 190 – 260 nm over a time period of 1000 minutes with a full scan every 4 minutes (reaction time progressing from blue to red) [concentration of BSA 0.0027 mM], **B.** Initial rates of reaction against substrate concentration for the NBSA vs aqueous [C-PK][S₇] reaction (each data point representing the average of a minimum of 3 repeats).

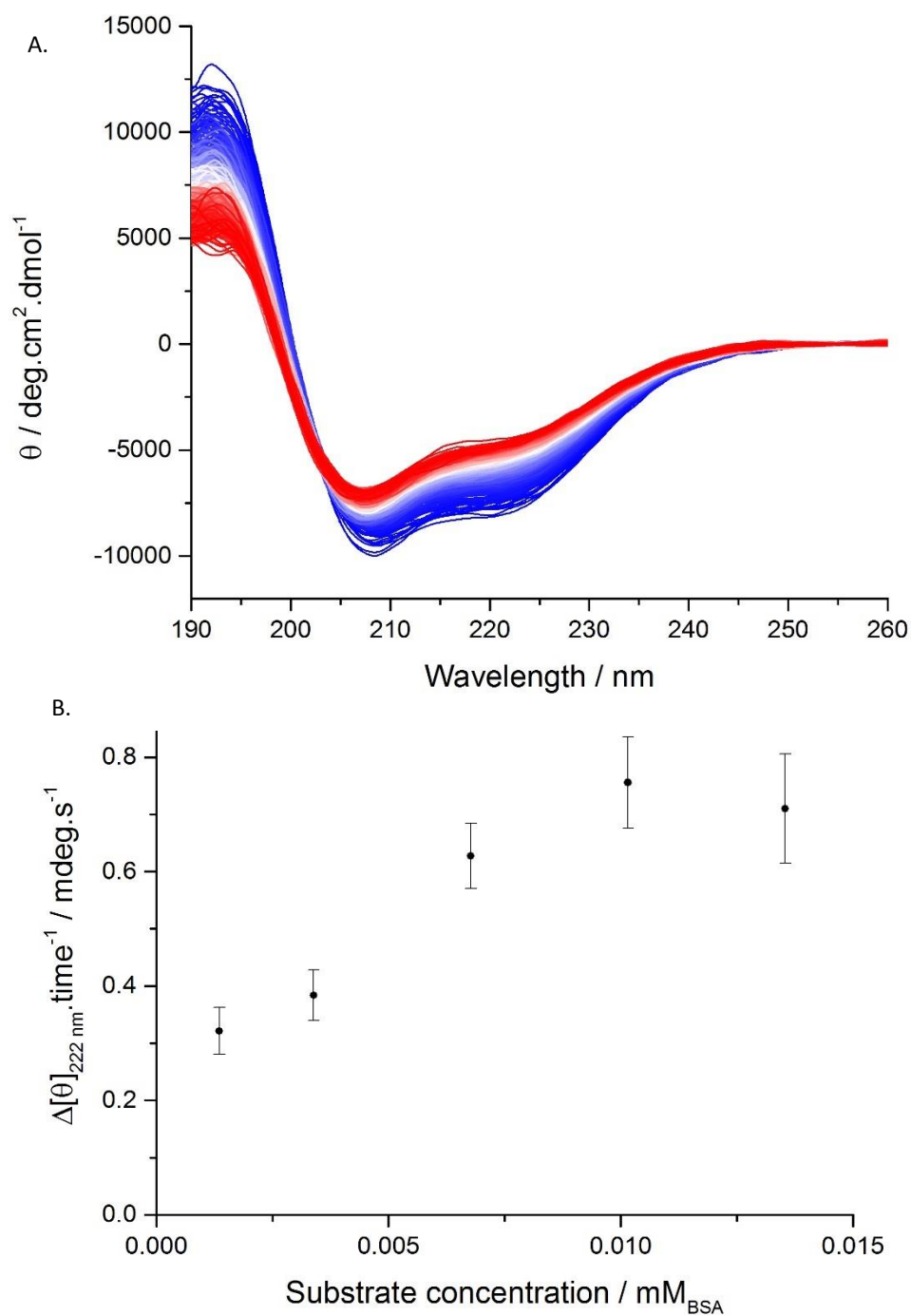


Figure 4.11: **A.** Mean residue ellipticity against wavelength showing far-UV CD spectra for the C-BSA vs aqueous [C-PK][S₇] reaction in the range of 190 – 260 nm over a time period of 1000 minutes with a full scan every 4 minutes (reaction time progressing from blue to red) [concentration of BSA 0.0027 mM], **B.** Initial rates of reaction against substrate concentration for the C-BSA vs aqueous [C-PK][S₇] reaction (each data point representing the average of a minimum of 3 repeats).

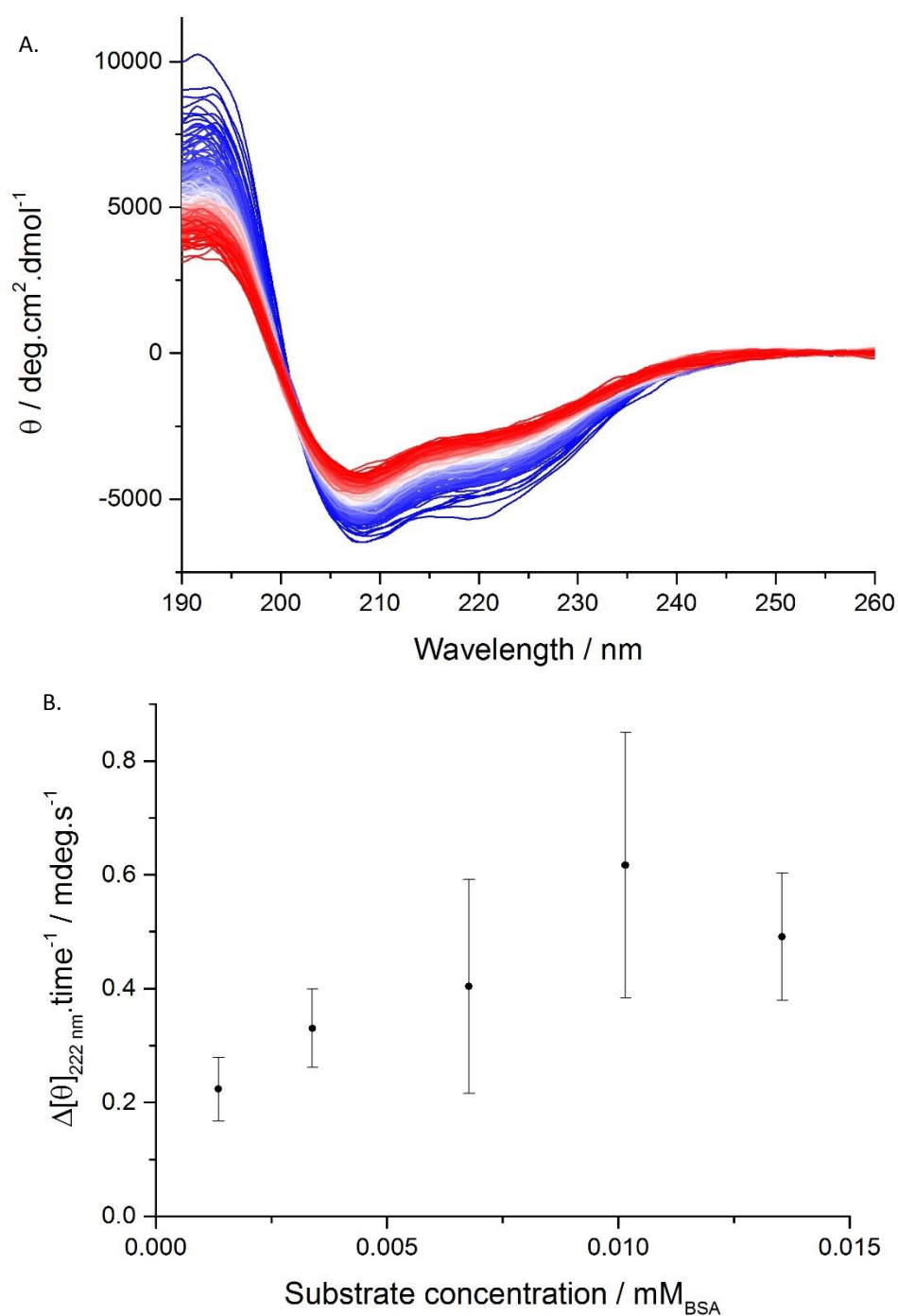


Figure 4.12: **A.** Mean residue ellipticity against wavelength showing far-UV CD spectra for the aqueous [C-BSA][S₇] vs aqueous [C-PK][S₇] reaction in the range of 190 – 260 nm over a time period of 4500 minutes with a full scan every 4 minutes (reaction time progressing from blue to red) [concentration of BSA 0.0027 mM], **B.** Initial rates of reaction against substrate concentration for the aqueous [C-BSA][S₇] vs aqueous [C-PK][S₇] reaction (each data point representing the average of a minimum of 3 repeats).

In comparison to NBSA, the surfactant conjugation of PK to form aqueous [C-PK][S₇] had a detrimental impact on the activity of the protease towards the substrate, with a 99% drop in V_{\max} relative to NBSA v NPK and a 95% drop in V_{\max} relative to NBSA v C-PK (Figure 4.10). In relation to C-BSA, the surfactant conjugation of PK to form aqueous [C-PK][S₇] had a detrimental impact on the activity of the protease towards the substrate, with a 79% drop in V_{\max} relative to C-BSA v C-PK and a 91% drop in V_{\max} relative to C-BSA v NPK (Figure 4.11). Relating to aqueous [C-BSA][S₇], the surfactant conjugation of PK to form aqueous [C-PK][S₇] had a detrimental impact on the activity of the protease towards the substrate, with a >99% drop in V_{\max} relative to aqueous [C-BSA][S₇] v NPK (Figure 4.12). These large decreases in V_{\max} observed for the pairings containing aqueous [C-PK][S₇] are in line with the prediction, with the two factors affecting the activity being the structural alteration and steric hindrance. It cannot be stated which factor affected the activity to a greater extent as the effects cannot be isolated. However, it is clear that the addition of surfactant to the surface of the protease has a negative impact on activity.

The degradation of N-BSA, C-BSA, and aqueous [C-BSA][S₇] using aqueous [C-PK][S₇] were all successful, albeit heavily retarded in relation to other pairings, and displayed evidence for the retention of protease activity once the protease had been surfactant conjugated, despite alteration to the secondary structure and high potential for steric hindrance.

4.3.1.2.3 Cationization of BSA

Cationization of BSA to form C-BSA was previously shown to change the secondary structure of the protein (Section 3.2.1.4.1), with a 19% decrease of α -helical content and a 12% increase in β -sheet content, a change that initiated aggregation of C-BSA forming large C-BSA clusters. This large conformational adjustment is likely to alter the externally displayed residues, implying that the initial preferential cleavage sites for the protease could differ, as cleaving peptide bonds that are generally not externally displayed before external peptide bonds may have a more rapid impact on the observed structure loss for pairings involving C-BSA. However, there was no clear change to the degradation pathway relative to the other pairings. The large increase of surface charge could have a two-fold detrimental impact on the activity. Firstly, being highly positively charged could increase any like charge repulsion between the protease and the substrate in the system, increasing the energy required for the protease to interact with the C-BSA and decreasing the probability of interaction. Secondly, having more cationic point charges on the surface could also disrupt the interaction of the binding cleft and the catalytic triad with C-BSA. The disruption of interactions and creation of aggregates, along with the known structural changes altering the displayed residues, makes it highly likely that the cationization of the substrate will alter the activity of each protease towards the C-BSA substrate, relative to N-PK vs N-BSA.

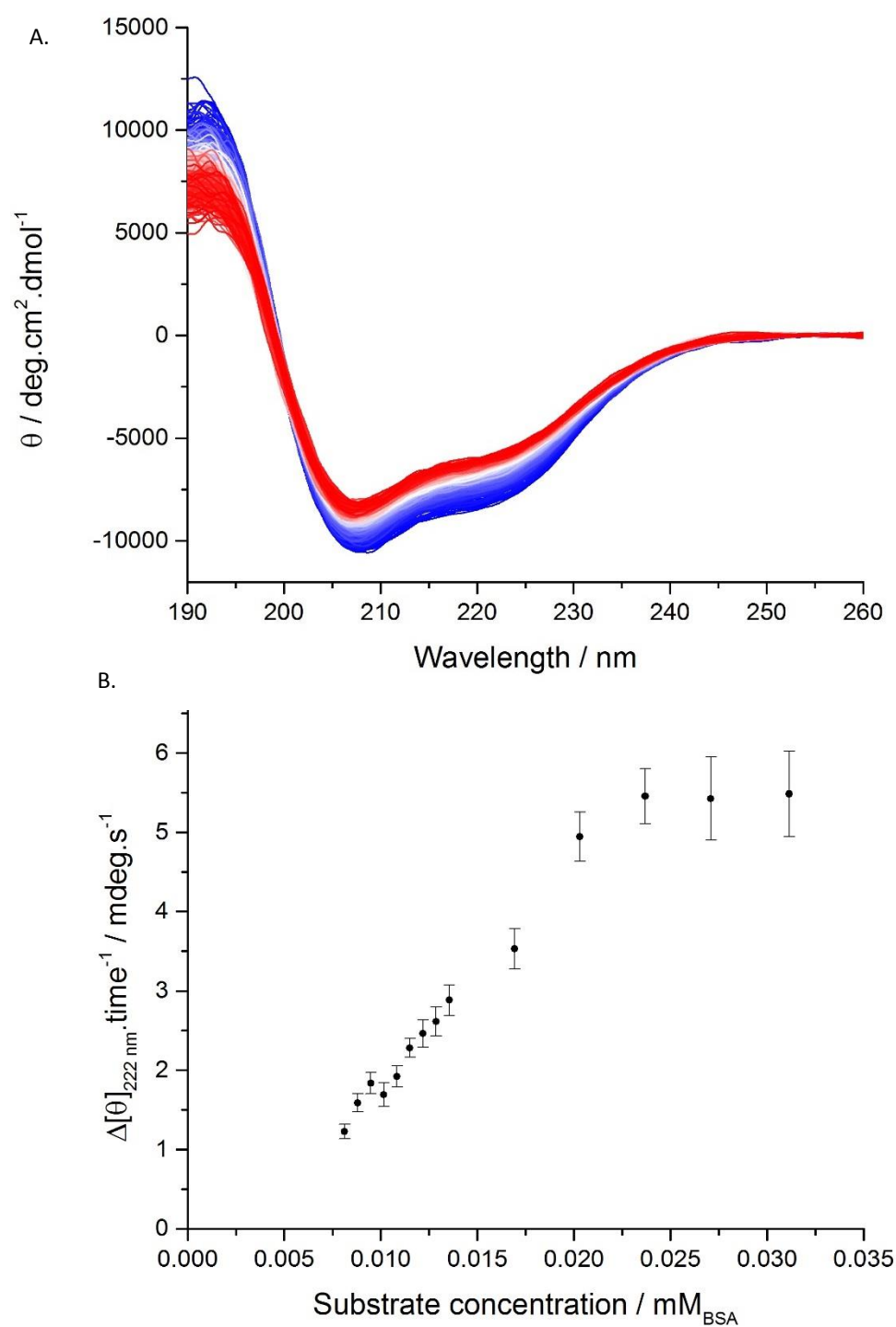


Figure 4.13: **A.** Mean residue ellipticity against wavelength showing far-UV CD spectra for the C-BSA vs NPK reaction in the range of 190 – 260 nm over a time period of 900 minutes with a full scan every 2 minutes (reaction time progressing from blue to red) [concentration of BSA 0.0027 mM], **B.** Initial rates of reaction against substrate concentration for the C-BSA vs NPK reaction (each data point representing the average of a minimum of 3 repeats).

In relation to NPK, the cationization of BSA to form C-BSA had a detrimental impact on the activity of the protease towards the substrate, with a 92% drop in V_{\max} relative to NBSA v NPK (Figure 4.6). This large decrease in activity was also seen in relation to C-PK, with a drop of 87% relative to NBSA v C-PK (Figure 4.7), and with aqueous [C-PK][S₇], with a drop of 40% relative to NBSA v aqueous [C-PK][S₇] (Figure 4.11). The cationization of BSA was shown to create aggregates of C-BSA. These aggregates, being highly charged, would require a longer time period to degrade due to the increase in repulsive forces between the aggregate and the protease. Interestingly, the degradation pathway of all C-BSA pairings matched those of the other pairings (discounting the outlier pairing), indicating that whilst the degradation was of an aggregate of C-BSA, the action of proteolytic degradation was the same, *i.e.* there was a similar cleavage pathway and unfolding once cleaved. Along with the larger size, the increase in positive charges on the surface that increases electrostatic repulsion between the substrate and the protease decreases the activity. This additional repulsion would increase the energy required for interaction, decreasing the probability of proteolytic cleavage; the effect being accentuated in the pairing with C-PK, where both surfaces are highly positively charged. The action of aqueous [C-PK][S₇] towards C-BSA displayed a low reduction in activity relative to the other C-BSA pairings, thought to be due to the surfactant corona dissipating the effect of the increased positive surface charge. This could therefore indicate that the decrease in activity is due to the presence of increased charge on the surface disrupting the interaction of the surface with the binding cleft within the active site.

Overall, the cationization of BSA has a detrimental impact on the activity of all proteases cleaving it, due to the formation of C-BSA aggregates, the increased number of cationic charges and the resulting higher surface charge. As the degradation pathways of each C-BSA pairing all look similar it is not thought that the alteration to the secondary structure of the BSA had a large effect on the observed protease activity. As the cationization of BSA had a lower impact on aqueous [C-PK][S₇] than other proteases it is hypothesised that the surfactant corona may have a repulsion dampening effect, reducing the hindrance of like charge repulsion.

4.3.1.2.4 Surfactant conjugation of C-BSA

Surfactant conjugation of BSA to form aqueous [C-BSA][S₇] was previously shown to marginally change the secondary structure of the protein relative to C-BSA (Section 3.2.2.3.2), with a 4% increase in β -sheet content linked to the dissolution of C-BSA aggregates; a change that could be linked to an increase the rate of degradation. The addition of anionic surfactant molecules to positively charged surface sites of C-BSA is thought to partially shield the many positive charges on the surface not only from each other but also from other proteins, reducing electrostatic like charge repulsion between the protease and the substrate; again, a change that would likely facilitate in an increase in the rate of degradation. As mentioned previously, the polymer chains are known to cover the surface of the protein¹⁰, therefore, when bound to the BSA, as when bound to the protease, through random walk

motion the tail sections can potentially enter the active site of proteases. This, again, would disrupt the interaction of the binding cleft with the protein surface. This steric hindrance makes it highly likely that the surfactant conjugation of BSA will reduce the activity of each protease towards aqueous [C-BSA][S₇] substrate, relative to NBSA and C-BSA pairings, despite the lack of aggregation and lower surface repulsion.

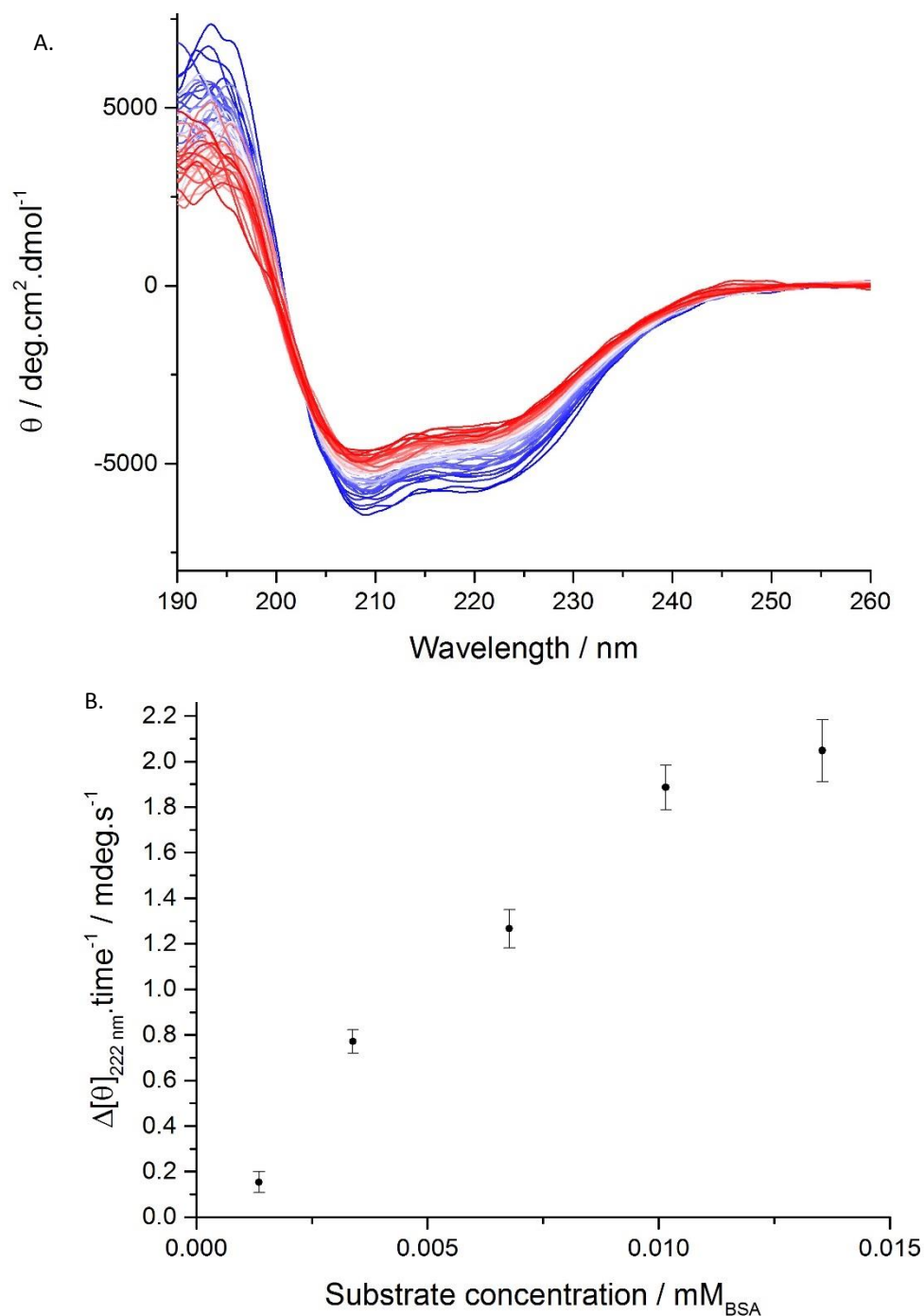


Figure 4.14: **A.** Mean residue ellipticity against wavelength showing far-UV CD spectra for the aqueous [C-BSA][S₇] vs NPK reaction in the range of 190 – 260 nm over a time period of 120 minutes with a full scan every 2 minutes (reaction time progressing from blue to red) [concentration of BSA 0.0027 mM], **B.** Initial rates of reaction against substrate concentration for the aqueous [C-BSA][S₇] vs NPK reaction (each data point representing the average of a minimum of 3 repeats).

In relation to NPK, the surfactant conjugation of C-BSA to form aqueous [C-BSA][S₇] had a negative impact on the activity of the protease towards the substrate, with a 43% drop in V_{\max} relative to C-BSA v NPK (Figure 4.13). This lower decrease in activity, relative to other pairings, was also seen in the comparison to aqueous [C-PK][S₇], with a 29% drop in V_{\max} relative to C-BSA v aqueous [C-BSA][S₇] (Figure 4.11). The decrease of the rate of proteolytic cleavage is attributed to the steric hindrance related to the surfactant molecules preventing interaction with the protease and potentially blocking the active site.

Overall, the surfactant conjugation of C-BSA has a negative impact on the activity of all proteases cleaving it, with the polymer chains reducing the probability of interaction and also potentially sterically hindering the active site of the protease. As the surfactant conjugation of C-BSA had a lower impact on aqueous [C-PK][S₇] than other proteases it is hypothesised that the surfactant coronas may be favourably interacting, marginally mitigating the negative influence of the steric hindrance, whilst still retarding the activity rate due to the large increase in surfactant molecules.

The mechanism of proteolytic action towards aqueous [C-BSA][S₇] remains unknown, and whilst the rate of activity for both pairings containing aqueous [C-BSA][S₇] was lower relative to N-BSA vs N-PK (more than 20-200 times slower), the observation of any proteolytic activity was an unexpected and interesting result due to the large increase in steric hindrance. One proposed mechanism is that the presence of the surfactant near to or conjugated to a cleavable residue (like a scissile bond) merely retards the proteolytic activity, rather than preventing it, however, this is highly unlikely due to the size of the surfactant relative to the protease active site. A second plausible mechanism is that the presence of the protease active site is a high enough energetic drive to replace the electrostatic attraction between the anionic surfactant and the positively charged surface site with the numerous hydrogen bonds that form once the binding cleft has interacted with the substrate. The now detached surfactant molecule could remain interacting with the rest of the coronal molecules, potentially re-binding once the protease has recycled the catalytic site, or instead remaining within the favourable corona. The third and most probable mechanism that is postulated is that the areas of surfactant conjugation on the surface of BSA were not protease active areas. That is to say, the areas that were left un-modified, or were far enough away from any surfactant binding sites, were lengths of the primary sequence that were cleavable by PK. It is known that proteases do not cleave areas of a proteins surface that are involved with folding, *i.e.* they don't bind to residues that are part of secondary structure motifs⁹. If all of these 'cleavable' areas of the substrate's surface were too close to or were electrostatically conjugated to the surfactant then it would be even more unlikely that the protease would be capable of binding, therefore, some cleavable sites must remain. However, further

investigation into the mechanistic dynamics of this complex system would be required to determine the mechanism.

4.3.1.3 K_m (Binding affinity)

The binding affinity (K_m) is calculated using the Michaelis-Menten equation (Section 2.1.3) and elucidates the strength of the binding interaction between the enzyme and the substrate. For these calculations the enzyme is either PK, C-PK, or aqueous [C-PK][S₇] and the substrate is either BSA, C-BSA, or aqueous [C-BSA][S₇]. The larger the calculated K_m value, the weaker the binding affinity and conversely, the smaller the calculated K_m value, the stronger the binding affinity.

Relative to N-BSA vs N-PK (0.059 mM), the cationization of BSA reduces the binding affinity and the cationization of PK increases the binding affinity (0.49 mM and 0.004 mM respectively). This implies that the highly cationized surface of C-BSA has weaker interactions with the binding cleft of PK than the unmodified surface due to the increase of charge. It is not thought that the prevalence of C-BSA aggregates in the system had an impact on the binding affinity. An increase in binding affinity for N-BSA vs C-PK indicates that a more energetically favourable interaction environment is generated once C-PK is formed. Interestingly, this indicates that the increase in α -helical content alters the active site environment of C-PK to be more favourable to substrate binding than native PK. These two effects are cancelled out in relation to C-BSA vs C-PK, with the binding affinity being similar in value to that of N-BSA vs N-PK (0.059 mM and 0.051 mM respectively). This return to a similar binding affinity demonstrates that the decrease in binding energy brought through the active site conformational alteration is overcome by the increase in interaction energy required to bind due to the increase in positive surface charge.

Interestingly, surfactant conjugation of C-BSA and C-PK decreased the value of K_m to a value lower than that of N-BSA vs N-PK (0.024 mM and 0.011 mM respectively, with a K_m of 0.003 mM for the combined system), indicating that the presence of surface-bound surfactant increases the binding affinity. This result was unexpected. It was hypothesised that the increase in binding affinity is rather due to an increase in steric hinderance in the substrate leaving the active site. As the binding affinity (K_d) relates the rate of enzyme binding (k^1) to enzyme de-binding (k^{-1}) (Equation 1), if de-binding occurs preferentially to binding, then the binding affinity will be higher. In the case of the protease/substrate pairing it is unlikely that the substrate protein or the cleaved protein products will bind more strongly to the protease active site. However, once the substrate has been cleaved a certain degree of unfolding would occur, potentially exposing hydrophobic residues that the alkyl tail domains of the surfactant could interact with. This additional interaction energy once cleaved would lead to the retardation of the removal of the products and increase the rate constant for de-binding and lower

the binding affinity. Therefore, the interaction between the protease and the substrate appears stronger. However, molecular modelling of these highly complex system would be required to provide evidence for this hypothesis.

$$K_d = \frac{k^1}{k^{-1}} \quad (1)$$

Equation 1: Where K_d is the binding affinity, k^1 is the rate constant for binding the substrate, and k^{-1} is the rate constant for de-binding the products¹².

4.4 Solvent-free liquid protease activity

4.4.1 Introduction

To investigate solvent-free liquid protease activity the enzyme assay developed for the aqueous precursor protease systems had to be altered, with two quartz plates being used rather than the quartz cuvettes as used for the aqueous systems. As the solvent-free proteins are only liquids at temperatures over 40 °C, assays were conducted at 45 °C rather than 25 °C, ensuring that both solvent-free proteins were melted. Both solvent-free liquids displayed different thermal stabilities, therefore, controls were required to confirm that any observed reduction in signal intensity was related to enzymatic activity rather than thermal degradation. Approximately equal masses of both solvent-free liquid [C-PK][S₇] and [C-BSA][S₇] were used for the controls as for the enzyme testing, allowing for ease of comparison. For the samples that were mixed, the correct masses of each protein were placed on two separate plates that were then sandwiched and thoroughly rubbed together ensuring maximal mixing.

The solvent-free liquid protease system brings about several challenges owing to the intrinsic material properties. The large increase in viscosity brought about by the removal of solvent from the system and due to vastly fewer available water molecules reducing the maximal enzyme turnover, it was hypothesised that the rate of enzyme activity would be lower if the ratio between BSA and PK was retained at 4:1 or higher (as with the aqueous samples). This is despite the large increase in concentration increasing the probability for protease substrate contact. The increase in viscosity would decrease the rate of diffusion within the system, slowing the movement of substrate into and products out of the active site. The reduction in available water molecules limits the maximum number of times PK can perform a proteolytic reaction, one water molecule being required for each reaction. Each solvent-free liquid [C-PK][S₇] and solvent-free liquid [C-BSA][S₇] construct has fewer than 30 water molecules (Section 3.2.3.3.2), therefore the maximum number of proteolytic reactions for each solvent-free liquid [C-PK][S₇] is fewer than 60 (30 from each construct), much lower than an aqueously solvated PK construct. Alongside this, the probability for one of these few water molecules entering the active site is low relative to the aqueously dispersed systems, again reducing the rate of activity.

As the activity was hypothesised to be lower, a larger percentage of the final protein concentration was solvent-free liquid [C-PK][S₇], the ratio being approximately 1:7 (BSA:PK). This is a large increase relative to the aqueously dispersed system (4:1 BSA:PK or higher in favour of BSA) and increases the probability of PK being the substrate protein. This was done to specifically increase the probability that an active solvent-free liquid [C-PK][S₇] construct that contains a water molecule within the active site would come into contact with a solvent-free liquid substrate (being either solvent-free liquid [C-PK][S₇] or [C-BSA][S₇]). An increase was required because it was assumed that a percentage of proteases are rendered inactive during the synthesis and through lack of active-site water molecules and as discussed previously, the rate of activity was hypothesised to be lower.

Therefore, if the drop in signal intensity observed for the mixed solvent-free system was larger than that of the combined controls, this would signify that the solvent-free liquid [C-PK][S₇] was active towards the substrate proteins (solvent-free liquid [C-PK][S₇] and solvent-free liquid [C-BSA][S₇]).

4.4.2 Solvent-free liquid protein activity results and discussion

The controls (Figure 4.15) show that being held in a 45 °C environment for 19 hrs did initiate some structural degradation in both the solvent-free liquid [C-PK][S₇] and the solvent-free liquid [C-BSA][S₇], with both solvent-free liquid proteins showing a similar reduction in 222 nm signal intensity (a drop of 15% each). Aqueously dispersed proteinase K has previously been shown to be relatively resistant to self-catalysis (PK cleaving PK), therefore, it was assumed that solvent-free liquid [C-PK][S₇] was also relatively resistant to self-catalysis. Furthermore, if there was a large amount of self-catalysis, a larger reduction in signal intensity would be expected. This would indicate that the reduction in signal for the solvent-free liquid [C-PK][S₇] was mainly due to thermal degradation. The two individual control spectra (that were taken in separate spectroscopic runs) can be combined to display the spectra as if both solvent-free liquid proteins were analysed simultaneously, but with no contact between the two liquids (Figure 4.16). This simulates a control for the setup of the enzymatic experiment where the two solvent-free liquid proteins are in contact (Figure 4.16). Therefore, this means that a reduction in signal intensity of approximately 15% is expected if solvent-free liquid [C-PK][S₇] is not active, and any appreciable further reduction in signal intensity is due to the solvent-free liquid [C-PK][S₇] retaining a level of activity towards solvent-free liquid [C-BSA][S₇].

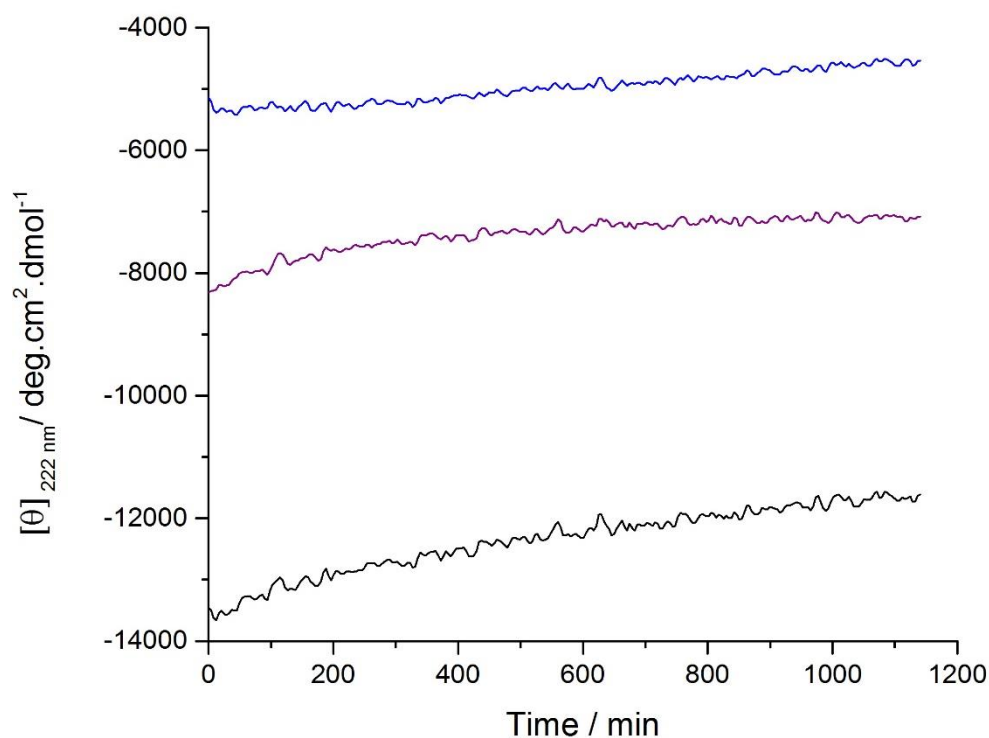


Figure 4.15: A. Mean residue ellipticity at 222 nm against time showing far-UV CD spectra for solvent-free liquid [C-PK][S₇] (blue), solvent-free liquid [C-BSA][S₇] (magenta), and the combined spectra of the two individual solvent-free liquid controls (black) at 45 °C.

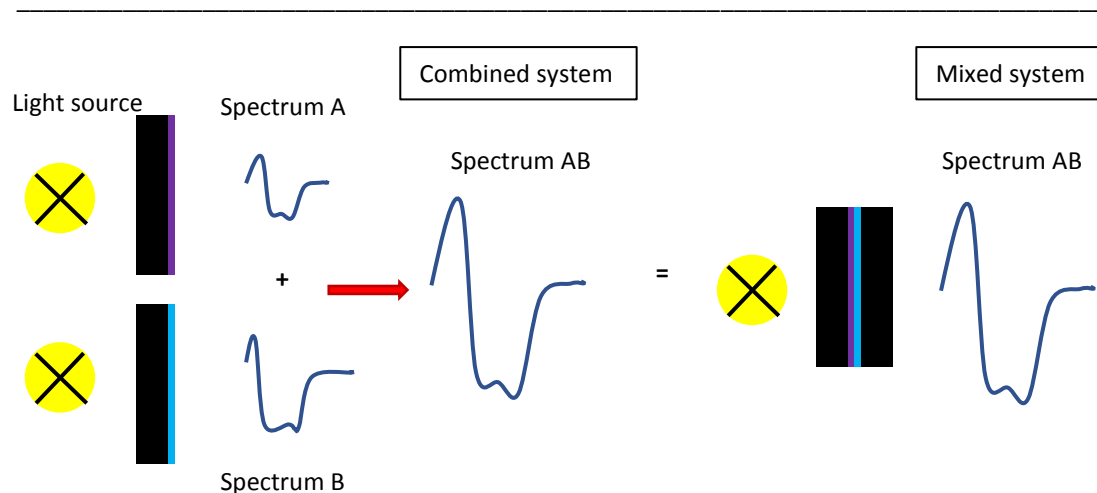


Figure 4.16: Graphic displaying the circular dichroism spectroscopy principle that two individual spectra of two different materials can be combined to give the same result as one spectrum taken of both samples. In this case, the two materials are solvent-free liquid [C-PK][S₇] (magenta) and [C-BSA][S₇] (blue) mounted on quartz plates (black). For illustrative purposes only one plate is shown for the individual melts, whereas two plates were used to sandwich each control experiment.

It is clear by comparing the control and the mixed solvent-free liquid system (Figure 4.17) that there is a substantially larger decrease in signal intensity when the two solvent-free liquid proteins are mixed together. As previously discussed, a reduction in signal intensity of more than the control level of 15% would indicate activity. The mixed solvent-free system reduced signal intensity by 45% over the same time. This, therefore, exhibits a degree of retained enzyme activity for solvent-free liquid [C-PK][S₇] towards a solvent-free liquid protein substrate. The persistence of enzymatic activity in the solvent-free liquid state brings about several questions: How long does the protease remain active before denaturation renders it inactive? How many times does the protease perform proteolysis? Is the mechanism for proteolysis in the solvent-free liquid state congruent with proteolysis in an aqueously dispersed system? Does the entropic stabilisation of the solvent-free liquid proteins alter how the cleaved protein unfolds?

Looking at the rate of signal reduction throughout the reaction time of the mixed system, it is evident that there are three main phases of conformational change. The first phase, from 0 – 200 min, has a very sharp reduction in signal intensity with the rate being approximately 4.5 times greater relative to the control (Table 4.2). In the second phase, from 200 – 400 min, the rate of signal reduction decreases whilst remaining approximately 2.5 times higher relative to the control rate (Table 4.2). The third phase, 400 – 1140 min, again sees a decline in the rate of signal reduction, however, the rate is now approximately equal to the control (Table 4.2). The decline in rate observed for the mixed sample indicates a loss of activity throughout the reaction period, an effect likely originating from inactivity due to the lack of water molecules within the active site or the protease being denatured. After 400 minutes the rates of the two systems become approximately equal, suggesting that all solvent-free liquid [C-PK][S₇] activity has ceased; the reduction in signal intensity being caused by thermal denaturation of the remaining secondary structure of both solvent-free liquid proteins. It should be noted that the rates cannot be directly compared to those of the aqueous precursors as those had consistent levels of protease concentration, levels that were many orders of magnitude lower than

that of the solvent-free system. However, it can be said with some certainty that the rate of proteolytic activity within the solvent-free system would be lower than that of all the aqueous systems because of the higher viscosity of the liquid, largely reduced water content, and the lower probability of water molecules entering the active site reducing the enzymatic turnover.

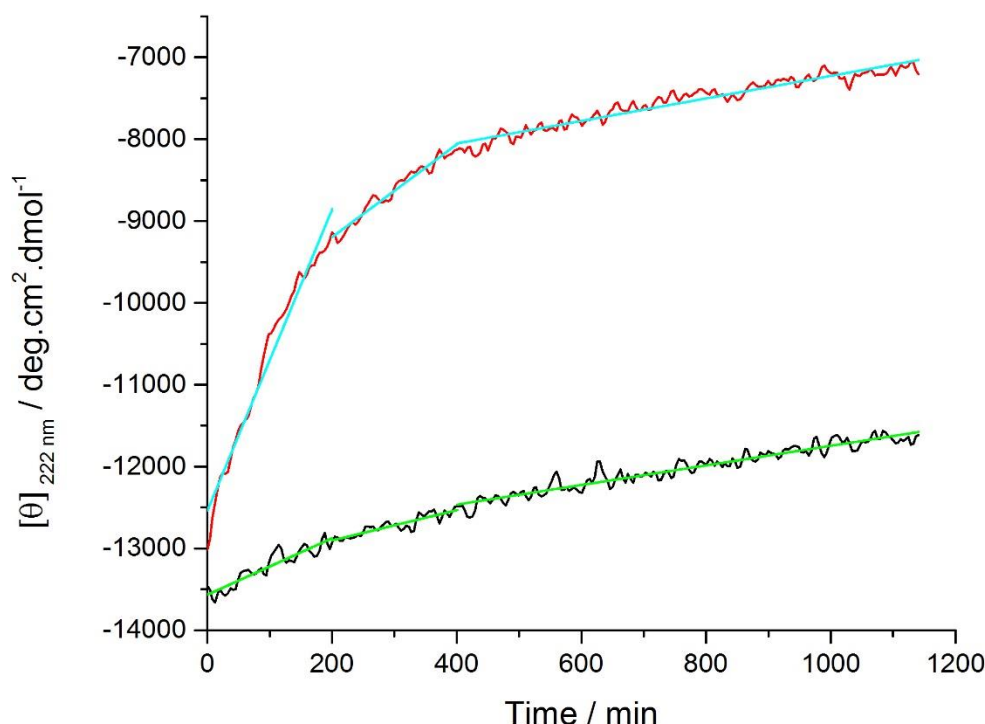


Figure 4.17: Mean residue ellipticity at 222 nm against time showing far-UV CD spectra for mixed solvent-free liquid [C-PK][S₇] and [C-BSA][S₇] (red) (fits for solvent-free rates in cyan) and the control (a combined spectrum of the two individual solvent-free liquid controls) (black) (fits for control rates in green) at 45 °C.

Table 4.2: Table displaying the rate of signal reduction of the mixed system and the combined system, along with the percentage increase in signal reduction.

Time period (min)	Rate of mixed sample (deg.cm ² .mol ⁻¹ .min ⁻¹)	Rate of control sample (deg.cm ² .mol ⁻¹ .min ⁻¹)	Percentage increase in signal reduction (%)
0 – 200	19.7	4.2	470
200 – 400	5.1	2.1	240
400 – 1140	1.2	1.1	110

However, this large difference in signal reduction between the control and the mixed sample does not indicate that all the additional signal reduction was directly related to the protease activity since a protein that has been cleaved is more likely to denature further and more rapidly than one that has not. Further to this, the unfolding of solvent-free liquid protein has the potential to expose hydrophobic core residues that will preferentially interact with the hydrophobic tail domains of the surfactant, increasing the denaturation of the protein. However, it is evident that the solvent-free liquid [C-PK][S₇] was initially active, with each individual construct cleaving substrate until rendered

inactive. Inactivity can arise in the solvent-free liquid [C-PK][S₇] by either no more water molecules entering the active site or by denaturation (caused by temperature and/or potential proteolytic cleavage). As the solvent-free liquid [C-PK][S₇] control has already shown, there was not a notable amount of self-catalytic cleavage since there was no drastic reduction in signal intensity and therefore the denaturation of the protease is mainly temperature based. This signifies that thermal denaturation of the protease occurs despite the increased entropic thermal stability, indicating that longer structural thermal studies are required to investigate the thermal behaviour of the solvent-free liquid protein over longer time periods.

Regarding the question of enzymatic turnover, from the data collected it cannot be categorically stated as to whether each solvent-free liquid [C-PK][S₇] construct performed one or multiple proteolytic reactions. However, the initial large reduction in signal intensity is unlikely to originate solely from one cleavage event on the majority of proteins (assuming that not every PK in the system diffuses towards and cleaves a BSA). This is because it is not likely that the cleavage of one external peptide bond of an aqueously solvated globular protein would cause a large unfolding event, therefore, because of the increased entropic barrier to unfolding brought about through the large increase in the concentration of the solvent-free liquid material, it is even less probable. Taking this into account, it is plausible that multiple solvent-free liquid [C-PK][S₇] constructs did perform more than one proteolytic reaction, indicating that enzymatic turnover was retained and the few available water molecules could diffuse into the active site.

As for the mechanism of proteolytic activity in the absence of solvent, this can presently only be speculated on, as with the mechanism for the proteolytic activity towards the surfactant conjugated substrate mechanism. As discussed earlier (Section 4.3.2.2.4), the observation of any proteolytic activity within these highly sterically hindered systems is unexpected, and increasingly unexpected for the solvent-free liquid protein system. Here, the three proposed mechanisms for the proteolytic activity are the same as previously stipulated; the first is that the presence of a surfactant molecule proximal to a scissile bond does not prevent the cleavage of that bond, the second being the presence of the protease binding cleft provides a more energetically favourable interaction with the substrate surface than the electrostatically bound surfactant molecule, and the third being that some surface areas that are proteolytically cleavable remain unmodified throughout the solvent-free liquid protein synthesis and are therefore available for binding and cleaving.

The first mechanism proposed remains unlikely despite the hydrophobic binding pocket due to the size relations between the binding cleft and the surfactant molecule. Now the protease/substrate pairing is in the solvent-free liquid state, the second proposed mechanism is less likely than when in the aqueous state. This is because the system is now so densely packed that the enthalpy gains for this to occur would have to overcome a large entropic barrier; it is likely that the protease binding to the surface leaving a detached surfactant is less entropically favourable than the binding of one surfactant molecule. However, the third proposed mechanism remains the most favourable. The protease

requires an extended length of non-folded residues for the binding cleft to interact with in order for the serine residue within the catalytic triad to cleave the scissile bond. Even though the solvent-free [C-BSA][S₇] substrate is now more folded than when aqueously solvated, indicating fewer or shorter chain lengths of non-folded residues, it is still in a molten globular state (Section 3.2.3.5.1). If the protease can access these unmodified surface areas of the BSA then it can bind and proteolytically cleave the bond. The steric hindrance from both surfactant coronas will retard the protease from interacting with the substrate protein, reducing the probability for binding but, as displayed above, there is strong evidence for proteolytic cleavage, albeit a slow process. As with the aqueously dispersed systems, further investigation through complex modelling or small angle neutron scattering would be required to provide evidence towards any of these hypotheses.

4.5 Conclusions and future work

A novel circular dichroism-based enzyme activity assay was developed to investigate how the surface modifications required to form a solvent-free liquid protein affected the proteolytic activity of a protease towards a protein-based substrate. This study on the effect of the specific surface modifications required to form a solvent-free liquid protein on the proteolytic ability of a protease towards a protein-based substrate was conducted using a novel circular dichroism-based enzyme assay. CD is a technique often used to study the secondary structure of a protein, with a protein losing secondary structure (and therefore signal intensity) when it is cleaved by a protease. The assay has been demonstrated to be versatile with both liquid and solid phase samples through using a cuvette or sandwich cell CD set up respectively. The intrinsic absorption from the secondary structure of the proteins is collected, allowing the possibility to track the activity of any protease/substrate pairing with the substrate also being a protein. Therefore, this method of kinetic analysis does not require the addition of any external molecules, such as fluorophores, to allow for kinetic assessment. As such, the activity was tracked by following the reduction in signal intensity over time, utilising Michaelis-Menten kinetics to calculate values that relate to both the maximum rate of degradation and the binding affinity of the substrate within the active site.

The cationization stage was shown to bring about a reduction in enzymatic rate which was attributed to the increased positive surface charge increasing the repulsion between the protease and the substrate and to alterations in protease conformation. Cationization of PK had a greater effect on the proteolytic rate than the cationization of BSA because once surface-modified, the secondary structure of PK changes (Section 3.2.1.4.2), moving away from the native conformation. This has the consequence of potentially altering the shape of the active site and, therefore, the substrate-binding ability of the protease. However, the retention of activity in the cationized state does show the flexibility of the active site, preserving the fundamental binding interactions required for proteolytic activity even after a large overall conformational change in the protein. The formation of C-BSA was shown to create aggregates of C-BSA (Section 3.2.1.3.1), with PK and C-PK retaining activity towards the large constructs. Having an increased surface charge on the substrate has the effect of changing the potential binding sites upon the surface of the substrate altering the potential interactions with the active site. This was seen with C-BSA having a lower binding affinity with PK relative to N-BSA. When C-PK was tested against C-BSA the rate was lower than when only one of the pair was cationized. This was due to the much larger increase of like charge repulsion between the two proteins reducing the probability of interaction between those two proteins.

Electrostatic grafting of the anionic surfactant-polymer to the positively charged surface sites was shown to decrease the enzymatic rate further, this being attributed to an increase in steric hindrance from the surfactant-polymer. Regarding aqueous $[C-PK][S_7]$, this reduction in rate occurred despite the reclamation of a more native-like conformation (Section 3.2.2.3.3). In fact, the modification to C-PK

had a greater effect on the rate relative to C-BSA, indicating that the presence of surfactant upon the surface of C-PK was more detrimental to activity than its presence upon C-BSA. The polymer hindered potential surface binding sites on the surface of the substrate and may also have entered the active site, hindering the binding to the protein and the proteolytic action of the catalytic triad. Aqueous [C-PK][S₇] v aqueous [C-BSA][S₇] was shown to have the lowest catalytic rate of all aqueous pairings due to more polymer being present increasing the probability that it will enter and block the active site. This decrease once the substrate was surface modified indicates a potential method of protein delivery, protecting the required protein from proteolytic attack. Relative to native and cationized species, the binding affinity of the conjugated proteins was higher, potentially indicating that the protease would preferentially bind to aqueous [C-BSA][S₇] than to N-BSA. However, it was postulated that this high binding affinity was rather due to the alkyl tail domains of the surfactant interacting with the partially unfolded products, retarding the release from the active site, increasing the apparent binding affinity.

Solvent-free liquid protease activity towards a solvent-free liquid protein-based substrate was displayed for the first time. This, therefore, provides evidence towards water molecules only being required for the proteolytic reaction, rather than the stabilization of the protease, with the interaction with the surfactant molecules replicating the conformationally stabilizing interactions. Whether the solvent-free liquid protease had any enzymatic turnover could not conclusively be stated. However, due to the large reduction in secondary structure observed, it is likely that each protease performed more than one proteolytic reaction as the high entropic barrier that is present for unfolding in the solvent-free liquid state would otherwise take a high energy to overcome with just one proteolytic cleavage. The proposed mechanism for solvent-free protease interaction with a solvent-free substrate involves some cleavable surface sites remaining unmodified throughout the synthesis and remaining accessible to the protease binding cleft through the high steric hindrance brought about by the surfactant molecules. Due to the fewer available water molecules reducing the maximum enzymatic turnover and the potential for a reduced rate because of the high viscosity of the liquid proteins¹³, the PK:BSA ratio was increased (in favour of PK) relative to the aqueous pairings to ensure any present activity was observable over a reasonable timescale.

As this is the first occurrence of solvent-free liquid protease activity and also the first use of CD spectroscopy to track enzyme activity, there is a large amount of future refining that could occur. This could involve alteration of the PK:BSA ratio, optimisation of the reaction time, and the potential to increase the temperature, decreasing the viscosity of the liquids and evaluating the protease activity at a high temperature (as was done with solvent-free liquid lipase¹⁴), along with complex molecular modelling and small angle neutron scattering investigations of the system to understand the

interaction between the two solvent-free liquids. Whilst expanding the knowledge of how proteases function, further work could also be extended into the practical uses for protease activity in this unique solvent-free liquid state of matter. Further, the use of proteases in arid environments is no longer an insurmountable hurdle, with the required water level 2 to 3 orders of magnitude lower for solvent-free liquid proteins.

4.6 References

1. Solary, E., Eymin, B., Droin, N. & Haugg, M. Proteases, proteolysis, and apoptosis. *Cell Biol. Toxicol.* **14**, 121–132 (1998).
2. Zhang, G., Holler, T. & Napper, A. Protease Assays. *Assay Guid. Man.* 1–14 (2012). at <http://www.ncbi.nlm.nih.gov/books/NBK92006/pdf/Bookshelf_NBK92006.pdf>
3. Ong, I. L. H. & Yang, K.-L. Recent developments in protease activity assays and sensors. *Analyst* **142**, 1867–1881 (2017).
4. Martin, S. F., Hattersley, N., Samuel, I. D. W., Hay, R. T. & Tatham, M. H. A fluorescence-resonance-energy-transfer-based protease activity assay and its use to monitor paralog-specific small ubiquitin-like modifier processing. *Anal. Biochem.* **363**, 83–90 (2007).
5. Cupp-Enyard, C. Sigma's Non-specific Protease Activity Assay - Casein as a Substrate. *J. Vis. Exp.* 4–5 (2008). doi:10.3791/899
6. Chiang, C. F., Okou, D. T., Griffin, T. B., Verret, C. R. & Williams, M. N. Green fluorescent protein rendered susceptible to proteolysis: positions for protease-sensitive insertions. *Arch. Biochem. Biophys.* **394**, 229–35 (2001).
7. Ji, T. H. & Urry, D. W. Correlation of light scattering and absorption flattening effects with distortions in the circular dichroism patterns of mitochondrial membrane fragments. *Biochem. Biophys. Res. Commun.* **34**, 404–411 (1969).
8. Wang, J., Dauter, M. & Dauter, Z. What can be done with a good crystal and an accurate beamline? *Acta Crystallogr., Sect. D* **62**, 1475–1483 (2006).
9. Madala, P. K., Tyndall, J. D. A., Nall, T. & Fairlie, D. P. Update 1 of: Proteases universally recognize beta strands in their active sites. *Chem. Rev.* **110**, PR1-PR31 (2010).
10. Gallat, F.-X. *et al.* A polymer surfactant corona dynamically replaces water in solvent-free protein liquids and ensures macromolecular flexibility and activity. *J. Am. Chem. Soc.* **134**, 13168–71 (2012).
11. Petrotchenko, E. V. *et al.* Use of Proteinase K Nonspecific Digestion for Selective and Comprehensive Identification of Interpeptide Cross-links: Application to Prion Proteins. *Mol. Cell. Proteomics* **11**, M111.013524 (2012).
12. Nelson, David L.; Cox, M. M. *Lehninger principles of biochemistry*. (2013).
13. Perriman, A. W., Cölfen, H., Hughes, R. W., Barrie, C. L. & Mann, S. Solvent-free protein liquids and liquid crystals. *Angew. Chemie - Int. Ed.* **48**, 6242–6246 (2009).

14. Brogan, A. P. S., Sharma, K. P., Perriman, A. W. & Mann, S. Enzyme activity in liquid lipase melts as a step towards solvent-free biology at 150 °C. *Nat. Commun.* **5**, 6058 (2014).

Chapter 5 Encapsulation of solvent-free

liquid proteins within colloidosomes

5.1 Colloidosome introduction

Encapsulation is a method used to confine a chemical system, keeping it separate from each other system it may be in or be placed within through the use of a physical barrier, permitting control over the incorporation of the selected molecules¹. The barrier is often im- or semi- permeable, depending on the requirement of the capsule, with numerous techniques being employed to create a barrier with the desired permeability. One such capsule is a colloidosome; a colloidally stabilized Pickering emulsion that is crosslinked forming a shell, encapsulating the desired phase. The following section outlines Pickering emulsions, colloidosome formation, and the use of colloidosomes in the formation of red blood cell models.

5.1.1 Pickering emulsion overview

The stabilization of emulsions is known to be a key area of research interest, however, until the early 2000s the use of amphiphilic surfactants as emulsion stabilizers dominated the field, with, in comparison, very few investigations into the use of solid particles to stabilize an emulsion²⁻⁵. Emulsions stabilized *via* solid colloidal particles are known as Pickering emulsions and offer additional properties along with a different mechanism of stabilization to emulsions stabilized *via* surfactants.

Pickering emulsions provide benefits such as superior stability, low toxicity, and stimuli-responsiveness, when compared to their classical surfactant counter parts. An additional benefit is that the particles don't require an amphiphilic nature to function, with the mechanism of interaction between the two phases being adsorption. When particles are dispersed within one phase of a pair of immiscible liquids and an emulsion is formed they adsorb to the liquid-liquid interface that is present. The drive for this to occur is energetic, it being more energetically favourable for the particle to be interacting with both phases than for the two phases to be interact. This can be explained further through the change in the Gibbs free energy of interaction (ΔG):

$$\Delta G = -\pi R^2 \gamma_{ow} (1 - \cos \theta_c)^2 \quad (1)$$

Equation 1: Where R is the particle radius, γ_{ow} is the energy per unit area of oil-water interface, and θ_c is the contact angle (Figure 5.1).

The contact angle describes the affinity of the particle for each phase. If $\theta_c < 90^\circ$ then the particle is classed as hydrophilic and would preferentially stabilize o/w emulsions, and if $\theta_c > 90^\circ$ then the particle is classed as hydrophobic and would preferentially stabilize w/o emulsions. The transition angle is 90° , with the particle sitting equally within the two phases if $\theta_c = 90^\circ$, and also possessing the maximum desorption energy.

Here we “ignore all kinetic and external force fields (gravitational, electrical, optical, and magnetic, along with the surface roughness, ionic interactions, dielectric effects, and van der Waals interactions”². Even with particles of very small radius (< 1 nm) the absorption free energy (ΔG) is larger than that of the thermal energy ($k_B T$), meaning that the energy of the solvent molecules is often not large enough to remove the particles from the interface. However, if the contact angle varies too much it becomes comparable to $k_B T$ and the particles will move into one of the phases instead of sitting at the interface. The change in energy here is related to the loss of interface area, which is related to the angle relative to the contact line, defined through Young’s equation:

$$\cos \theta_c = \frac{\gamma_{po} - \gamma_{pw}}{\gamma_{ow}} \quad (2)$$

Equation 2: Where, γ_{po} is the energy per unit area of particle-oil interface and γ_{pw} is the energy per unit area of particle-water interface.

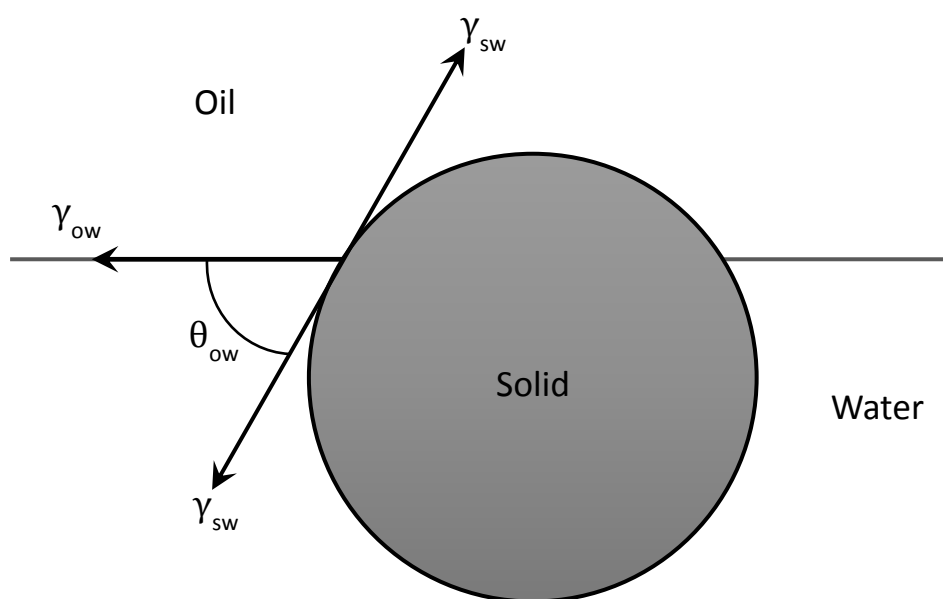


Figure 5.1: Diagram displaying a solid spherical particle at an oil/water interface and the various interfacial energies (oil/water (γ_{ow}), solid/water (γ_{sw}), and solid/oil (γ_{so})) and the oil/water contact angle created (θ_{ow}). All interfacial energies and contact angles are measured into the water phase (Diagram adapted from Binks²).

As the particle must be wetted by both the water and oil phases for it to be energetically favourable to sit at the interface, the surface displayed must contain both hydrophilic and hydrophobic moieties, the ratio of which will determine the contact angle. If it doesn't, the particle will remain dispersed within the favoured phase and the emulsion will be broken over time. If a particle were to display all the hydrophobic moieties on one half of the sphere and the hydrophilic on the other a Janus particle

would be created, and thus the particle will arrange in a specific orientation at the interface allowing for precise control of contact angle if the percentages of coverage were adjusted.

One of the most popular colloidal particles used for the stabilization of Pickering emulsions is silica, however, non-surface modified silica is unable to stabilize an emulsion due to the total wetting by water of silica. The surface is very easily hydrophobized *via* the surface-grafting of organosilanes, the final hydrophobic character being well controlled and selective. The formation of a solid casing through the crosslinking of the colloidal shell creates a colloidosome, named as such to be an analogy to liposomes, capsules formed through the use of characteristic phospholipid bilayers.

5.1.2 Colloidosomes

A colloidosome is a self-assembled cross-linked colloidal capsule that is formed *via* the colloidal-particle stabilization of an emulsion that is subsequently cross-linked to form a closely packed layer of colloidal particles that form a solid shell surrounding one or more components of the emulsion⁶. The packing of the shell is easily adjusted and can offer a high level of control over the permeability with the particle material and method for crosslinking offering control of the elasticity. The colloidal-particles used to stabilize the emulsion are determined by the hydrophobicity/hydrophilicity of the emulsion component that is to be encapsulated, how large the colloidosome is required to be post-templation, and the method utilised to crosslink the particles. Colloidosomes can be used to encapsulate a large variety of particles (organic, inorganic, and bio- molecules) gases, and liquids for many uses such as temporary or permanent storage, controlled release, and self-contained bioreactions.

5.1.2.1 Formation of colloidosomes

Colloidosomes are fabricated through a 3-step process. Firstly, the material that is to be encapsulated is emulsified within an immiscible liquid that has dispersed within it colloidal particles⁷⁻¹². Secondly, *via* self-assembly, the colloidal particles spontaneously adsorb onto the surface of the emulsion droplets, interacting with the interface between the two immiscible materials, forming a closed-packed layer of particles with a specified pore size (*i.e.* a Pickering emulsion). Thirdly, the layer of colloidal particles is crosslinked together in a controlled manner allowing for accurate control over the permeability through the pore size beneath; Fourthly, if required, the colloidosomes can be transferred from the 'emulsion solution' and placed within a solvent, typically one the same as the internal phase of the colloidosome, removing the interface stress by removing surface tension. This means that the permeability is fully controlled by the pore size and not influenced by the interfacial energies.

There are numerous techniques for the third step in this process, *e.g.* sintering, electrostatic binding, and chemical crosslinking, each being selected for the specific requirements of the colloidosome. Chemical crosslinking of the colloidal surfaces potentially gives a high level of control over both the thickness of the formed shell and the externally displayed functional groups. One such form of chemical crosslinking is silicification, the hydrolysis and condensation of organosilanes to the surface, forming a permeable oligomeric silica shell around the colloidal particles. Often with this method of crosslinking the colloidal particles chosen are silica particles, covalently binding the shell to the surface of the particles.

5.1.2.2 Silicification

Common silica-based molecules used for silicification are tetramethylorthosilicate (TMOS), $\text{Si}(\text{OCH}_3)_4$, or tetraethylorthosilicate, (TEOS), $\text{Si}(\text{OC}_2\text{H}_5)_4$, with TMOS being the preferred silicon alkoxide in systems where the reaction is required to be rapid, the shorter alkyl chain speeding up the initial stage of the reaction¹³. There are two stages of silicification: the initial stage is hydrolysis of one alkyl group forming a Si-OH bond (Figure 5.2), the second stage being the condensation of two of these groups forming an Si-O-Si bond, which can either be an alcohol condensation or a water condensation reaction (Figure 5.3).

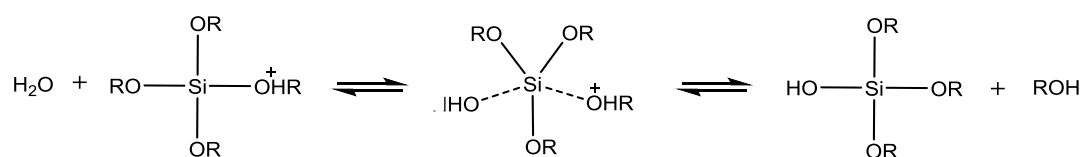


Figure 5.2: The mechanism of the acid-catalysed hydrolysis of a silicon alkoxide.

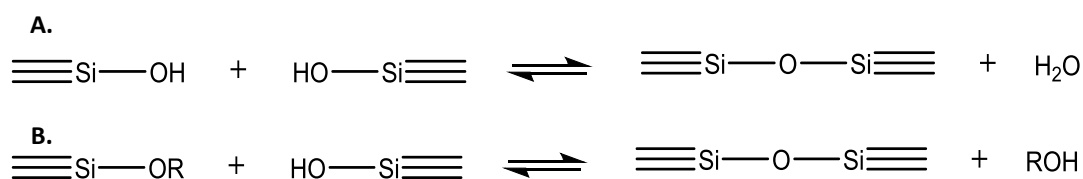


Figure 5.3: General mechanisms for the **A.** water condensation and **B.** alcohol condensation, of a partially hydrolysed silanol.

As the initial hydrolysis reaction only occurs in the presence of water, the silicon alkoxide is added to the oil phase of the emulsion. Once it reaches the interface with water where the silicon colloidal particles reside the reaction commences, with high potential for polycondensation reactions occurring between hydrolysed silanols and the silica surface. The thickness of the colloidosome's shell is determined by the ratio of colloid surface area to volume of added silicon alkoxide, with thin silica shells being semi-permeable.

5.1.3 Colloidosomes as an encapsulation technique

Pickering emulsions can be used to encapsulate a large variety of molecules/particles, with the formation of colloidosomes allowing for molecules to exist as distinct entities that are delineated from the surrounding phase, *e.g.* enzymes encapsulated within colloidosomes can function as self-contained bioreactors⁶. To encapsulate a molecule within a colloidosome, it requires dispersion within a suitable solvent, that is then formed into a Pickering emulsion with an immiscible solvent phase containing colloidal particles and subsequently cross-linked into a colloidosome. This colloidosome, containing the encapsulated species, can then be transferred into numerous alternative solvents, with the species being retained by the crosslinked colloidal-particle membrane. As the membrane can be permeable, with the permeability being controlled, selective substrate molecules can enter, and products can leave *via* diffusion across the membrane.

As colloidosomes can mimic primitive membrane-bound water microdroplets containing biochemical reactions, they have been widely studied as an inorganic form of a protocell (compartments that were present before the multi-compartmented cell evolved). Colloidosomes have been shown to replicate many different cellular functions such as: growth and division⁹, stimuli-responsive membrane-gated behaviour¹⁰, gene expression¹², and enzyme catalysis¹².

5.1.3.1 Synthetic red blood cell

One simple cell type to mimic is the red blood cell (RBC), being the method of oxygen (O_2) and carbon dioxide (CO_2) transportation around the body *via* the circulatory system. It is known to be lacking a nucleus and numerous organelles commonly found within other cells post maturation. The ejection of these, often thought vital, parts of a cell during evolution was to maximise the internal volume taken up by haemoglobin and also to allow the RBC to pass through constricting valves and capillaries along its route through the body. The 3-dimensional shape of the RBC, a biconcave disk possessing a dumbbell-shaped cross section (Figure 5.4), is such to optimise the flow properties within blood vessels and to increase the surface-area-to-volume ratio to increase the gas interactions with the surface.

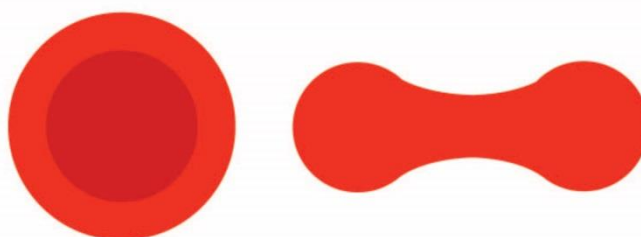


Figure 5.4: Diagram and cross-section of a red blood cell.

The fabrication of a synthetic red blood cell (sRBC) has several associated challenges to be overcome such as: formation of the characteristic RBC flexible morphology, the creation of a reliable, reusable, and stable source of viable haemoglobin, and the attainment of the high concentration of haemoglobin found within a RBC. 'Synthetic' blood has already been created from stem cell differentiation to RBCs, however, this is many decades from being a viable substitute due to issues such as blood typing that need to be overcome. A number of RBC models that mimic the complex morphological properties of a RBC have been synthesised¹⁴ and artificial haemoglobin that is stable for over 12 months with a circulatory life time of up to 36 hour has previously been created¹⁵.

This, therefore, demonstrates that two out of the three main challenges for the creation of sRBCs have been accomplished, with the creation of flexible biconcave discoidal capsules and of stable 'synthetic' haemoglobin. However, the issue of creating a sRBC that contains the high level of haemoglobin that natural RBCs possess has yet to be accomplished. Haemoglobin can reach concentrations of approximately 320 mg.mL⁻¹ within a RBC, attained by a complex mixture of various buffers, and is currently unattainable when attempting to fabricate a sRBC¹⁶. This is partially due to the solubility of lyophilized haemoglobin being only 20 mg.mL⁻¹ in aqueous un-buffered solutions, 16 times lower than within the natural RBC.

Solvent-free liquid proteins have displayed the ability to exist at high concentrations, as high as 400 mg.mL⁻¹, in a biocompatible-surfactant-conjugated protein nanocomposite state. If haemoglobin could be surface modified to form a solvent-free liquid haemoglobin, subsequently be encapsulated within a protocell model (colloidosome), and exhibit oxygen binding/de-binding, then this could be a method to combat the current high concentration issue seen with sRBC and see the creation of a novel proto-blood cell. Along with the high concentrations displayed with solvent-free liquid proteins, there is also known structural retention, signifying that the haemoglobin tetramer and the ability to bind/de-bind oxygen should be retained during the process. If the encapsulation within colloidosomes is successful, then an additional stage would be to encapsulate the solvent-free liquid haemoglobin within one of the morphologically correct sRBCs. This would effectively combine the three challenges to making a sRBC: formation of the characteristic RBC flexible morphology, the creation of a reliable, reusable, and stable source of viable haemoglobin, and the attainment of the high concentration levels of haemoglobin found within a RBC.

5.1.3.1.1 Encapsulation of haemoglobin

For the encapsulation to be successful the solvent-free liquid haemoglobin could be dispersed within a minimal volume of the aqueous phase of an emulsion that contains colloidal particles, *e.g.* silica nanoparticles. The w/o emulsion would therefore be stabilized and can be cross-linked, *e.g. via* silicification, and can subsequently be transferred to water. An additional option also exists that could increase the concentration further. Rather than dispersing the solvent-free liquid haemoglobin within the aqueous phase, it could instead be the opposing phase of the emulsion. As the externally displayed surfactant is tuneable, one could be selected that does not dissolve within the oil and would instead

sit at the interface with the oil phase containing colloidal particles. The colloidal particles would also sit at the interface, again allowing them to be successfully crosslinked and transferred to water.

Another oxygen carrying protein that has previously been modified to form a solvent-free protein liquid is myoglobin. Myoglobin displayed a high retention of native secondary and tertiary structure along with the ability to bind and de-bind several gases in this solvent-free state. Solvent-free liquid myoglobin can therefore be used as a substitute for haemoglobin during preliminary testing for the encapsulation of a solvent-free liquid protein and subsequent oxygen binding testing. If displayed to be successful, a solvent-free liquid haemoglobin can be created, and further testing can be carried out.

In this chapter myoglobin is taken through the solvent-free liquid protein synthesis and is characterised after each stage of the synthesis to confirm the work conducted by Brogan *et al*^{17,18}. Then, utilising the high protein density of the solvent-free liquid myoglobin, a high density of surface modified myoglobin nanoconjugate is encapsulated within a nano-clay stabilized colloidosome. These colloidosomes are then characterised and the location of the myoglobin nanoconjugate is investigated. This work forms the preliminary testing for future work involving the encapsulation of surface modified haemoglobin nanoconjugate, leading to the generation of a new synthetic red blood cell.

5.2 Results: Characterisation of solvent-free liquid myoglobin and aqueous precursors

The solvent-free liquid protein synthesis (detailed in Section 2.3.2.1) was carried out on the protein myoglobin, with each stage being characterised to assess how the synthesis altered myoglobin. In the following characterisation sections C-Myo refers to cationized myoglobin, aqueous [C-Myo][S] refers to surfactant conjugated myoglobin, and solvent-free liquid [C-Myo][S] refers to solvent-free liquid myoglobin.

5.2.1 Mass spectrometry

For native myoglobin the molecular mass was calculated to be 16992 g.mol^{-1} , as displayed by a sharp peak in the MS spectrum (Figure 5.5), fitting with literature values¹⁹. Within this spectrum, there was also a secondary shoulder at a lower intensity (≈ 0.2) thought to occur from the myoglobin interacting with an impurity in the sample. The number of acidic solvent-accessible surface residues present on myoglobin was calculated to be 21 (13 glutamic acid residues and 8 aspartate residues)¹⁹. Previous cationization work conducted on myoglobin found the optimum cationization pH to be pH 5.5, so this pH value was used. The cationization efficiency obtained was 71% (15 DMAPA molecules), lower than the 87% obtained in literature but sufficient for the formation of solvent-free liquid myoglobin.

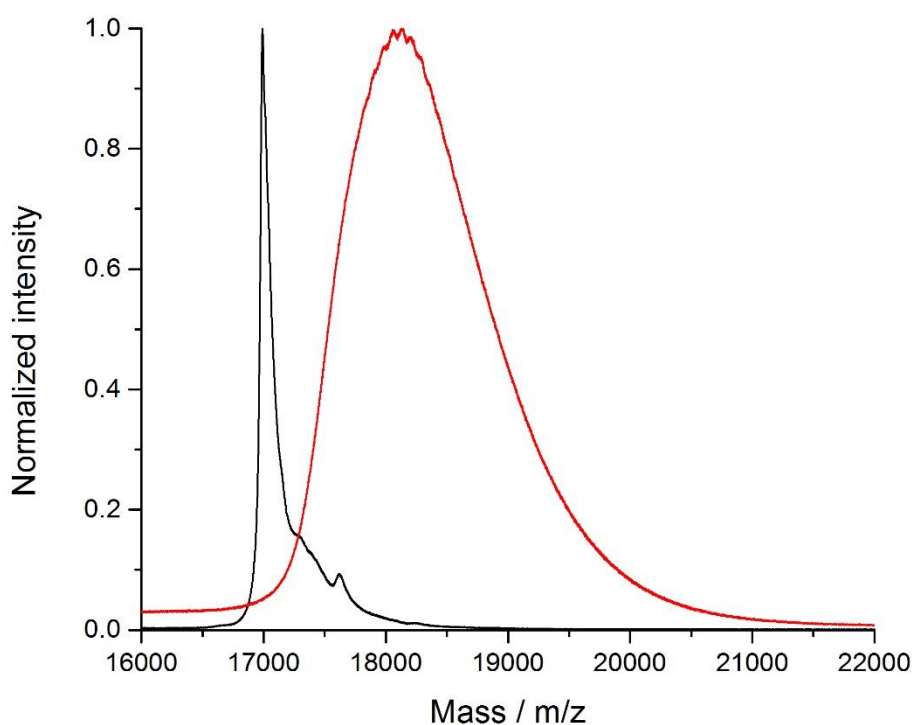


Figure 5.5: MALDI-TOF MS spectra for native myoglobin (black), and cationized myoglobin (red).

5.2.2 Charge density measurements using zeta potential

The zeta potential measurements of Myo, C-Myo, and aqueous [C-Myo][S], conducted in unbuffered water pH 6 - 6.5, yielded ZP values of $\xi = 18 \pm 2$ mV, 2 ± 1 mV, and -5 ± 1 mV respectively. These results provide further evidence for the successful cationization of Mb, with a charge density increase once cationized. The exhibited decrease in charge density once surfactant conjugated shows that there were charge shielding effects occurring once conjugated.

5.2.3 Particle size and distribution through dynamic light scattering (DLS)

The calculated results from the number particle distribution data for Myo and CMyo were 1.5 nm and 1.7 nm (Figure 5.6) respectively, both values fitting with values in literature¹⁷. The increase in size upon cationization corresponds to the length of the additional amines on the surface of Myo, known to be less than 0.1 nm in length. Again, as with BSA and PK, the justification for the use of number particle size distribution comes from the comparing of the number particle size distribution data and the intensity correlation data with previous literature values.

The calculated result from the number particle distribution data for aqueous [C-Myo][S₇] was 2.3 nm (Figure 5.6), again fitting with values in literature¹⁷. This increase displayed a similar increase in radius when compared to the increases observed when forming aqueous [C-BSA][S₇] and aqueous [C-PK][S₇], it being predominantly caused by the size of the surfactant molecules with the potential for a small increase in size due to conformational rearrangement. Alongside the average size increase the distribution of sizes also increases potentially displaying the formation of multiple coronal layers or a small amount of aggregation. However, their potential aggregates can be no more than a few nano-conjugates in number as the maximum size obtained is ≈ 12 nm.

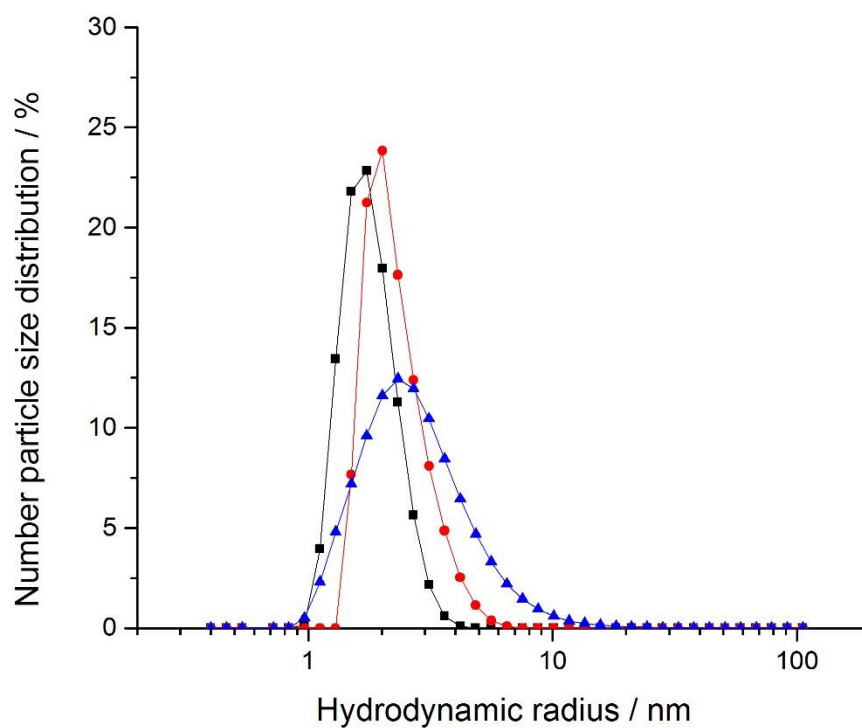


Figure 5.6: The number particle size distribution for aqueous Myo (black), aqueous C-Myo (red), and aqueous [C-Myo][S₇] (blue). [Conditions: unbuffered solutions at pH values of 6-6.5, with proteins concentrations of 0.5 mg.mL⁻¹, 25 °C]. Data displayed is an accumulation of at least 10 scans.

5.2.4 Differential scanning calorimetry

5.2.4.1 DSC of solvent-free liquid [C-Myo][S₁]

DSC traces of solvent-free liquid [C-Myo][S₁] displayed a sharp reversible exothermic crystallization peak at 0 °C in the heating cycle and sharp reversible endothermic melting transition peak at 22.5 °C in the heating cycle, with a small shoulder at 29 °C (Figure 5.7). These peaks were accompanied by observable glass transitions occurring at -51 °C in the heating cycle and at -49 °C in the cooling cycle. These thermal transitions contrast with the ones displayed on the DSC traces for S₁, both in value and shape. S₁ displayed a sharp reversible exothermic crystallization transition peak at -16 °C and sharp reversible endothermic melting transition peak at 32 °C in the heating cycle. These peaks were accompanied by an observable glass transition occurring at -47 °C in the heating cycle, a small broad exothermic transition peak at -20 °C and at second glass transition at -27 °C in the cooling cycle. For solvent-free liquid [C-Myo][S₁], the crystallization temperature was 16.3 °C higher and the melting transition temperature was 10.5 °C lower than that of S₁, along with a slight shift in glass transition temperatures. It is also observable that the transitions observed in S₁ have a higher heat flow change.

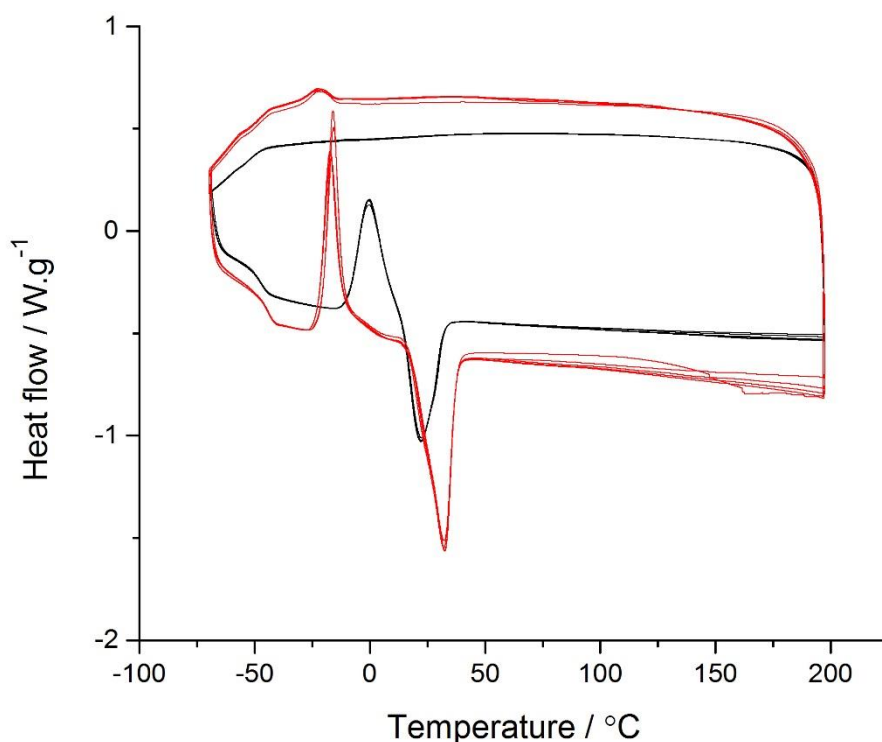


Figure 5.7: Cyclic DSC traces showing normalized heat flow against temperature for solvent-free liquid [C-Myo][S₁] (black) and S₁ (red). [Experiments were cycled a minimum of 6 times and the scan rate was 10 °C.min⁻¹].

5.2.4.2 DSC of solvent-free liquid [C-Myo][S₇]

DSC traces of solvent-free liquid [C-Myo][S₇] displayed a sharp exothermic crystallization peak at 1.9 °C in the cooling cycle that displayed slight hysteresis and sharp reversible endothermic melt transition peak at 44.2 °C in the heating cycle, with a small shoulder at 38 °C (Figure 5.8). These thermal transitions contrast with the ones displayed on the DSC traces for S₇, both in value and shape. S₇ displayed a sharp reversible exothermic crystallization transition peak at -1 °C and sharp reversible endothermic melting transition peak at 40 °C in the heating cycle. For solvent-free liquid [C-Myo][S₇], the crystallization temperature was 2.9 °C higher and the melting transition temperature was 4.2 °C higher than that of S₇. It is also observable that the transitions observed in S₇ have a higher heat flow change.

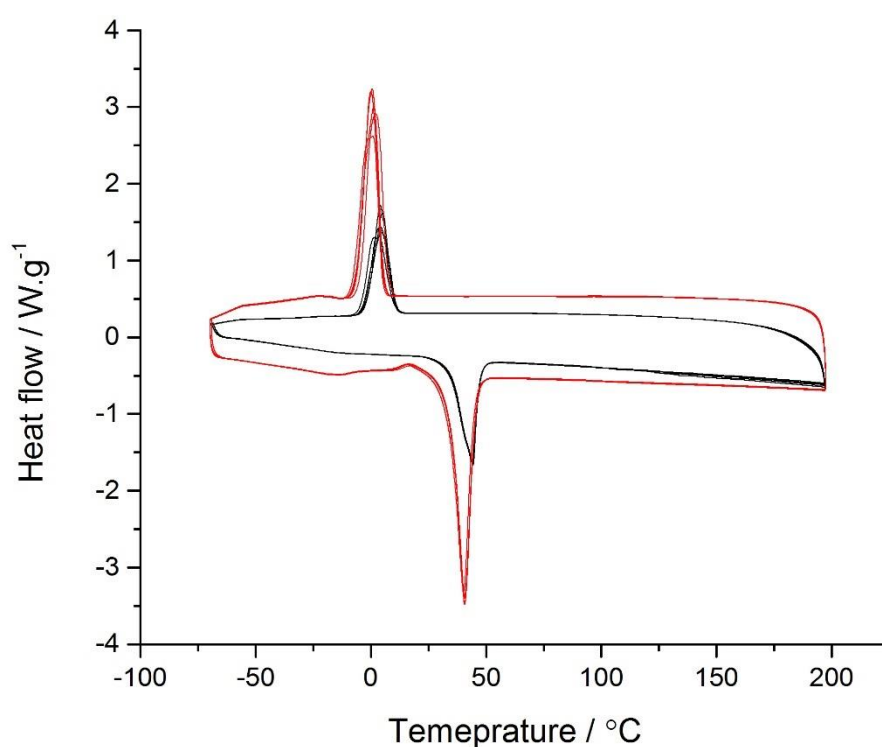


Figure 5.8: Cyclic DSC traces showing normalized heat flow against temperature for solvent-free liquid [C-Myo][S₇] (black) and S₇ (red). [Experiments were cycled a minimum of 6 times and the scan rate was 10 °C.min⁻¹].

The difference in thermal properties, *i.e.* the shift in melting and crystallization transition temperatures, between S₁ and solvent-free liquid [C-Myo][S₁], and between S₇ and solvent-free liquid [C-Myo][S₇], demonstrate that the solvent-free liquid proteins are distinct materials from the surfactants formed of discrete particles, and not only Mb that has been dissolved within a surfactant.

5.2.5 Thermogravimetric analysis

5.2.5.1 TGA of solvent-free liquid [C-Myo][S₁]

TGA of solvent-free liquid [C-Myo][S₁] displayed a decomposition temperature of 400 °C and a water content of 0.39% (Figure 5.9). The thermal decomposition temperature of lyophilized Mb is 345 °C, displaying an increased thermal stability of the solvent-free liquid [C-Myo][S₁] construct of 55 °C, again being comparable to the BSA analogue. The water content of 0.15 % for solvent-free liquid [C-Myo][S₁] was calculated; the total mass of a solvent-free liquid [C-Myo][S₁] construct was calculated to be 63198 g.mol⁻¹; the mass of C-Myo was 20106 g.mol⁻¹ and the mass of 36 S₁ surfactant molecules being 43092 g.mol⁻¹, yielding 5 water molecules that are highly associated with the protein construct.

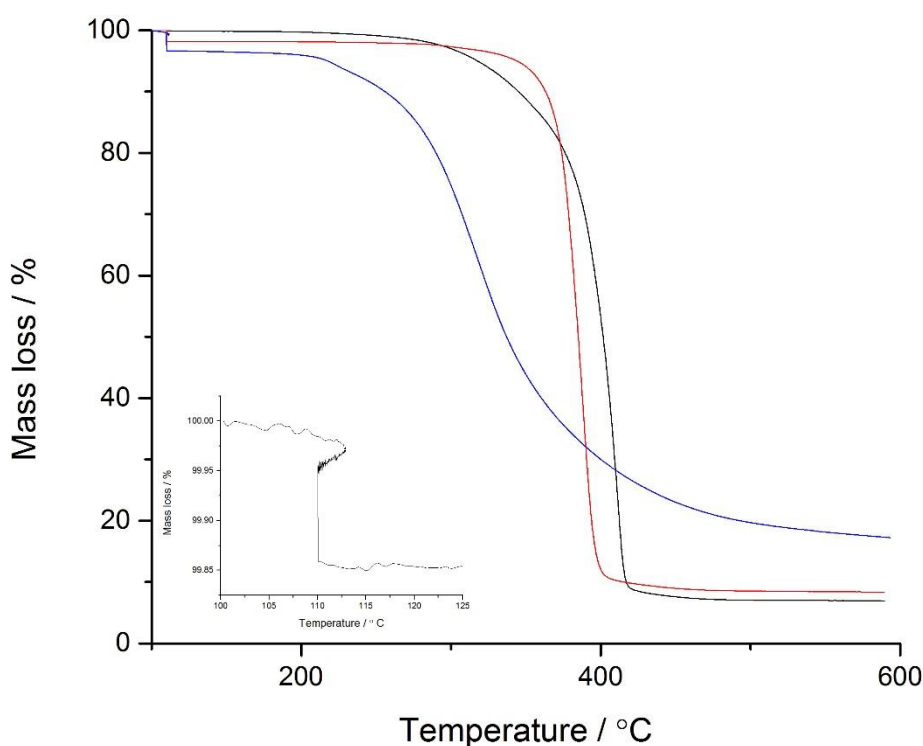


Figure 5.9: Thermogravimetric analysis of solvent-free liquid [C-Myo][S₁] (black), lyophilized Mb (blue), and S₁ surfactant (red), showing percentage mass loss in relation to temperature. The inset displays a magnified region of solvent-free liquid [C-Myo][S₁] in the range 100 - 120 °C showing the mass loss associated with the loss of surface-bound water (temperature held at 110 °C for 1 hour), *i.e.* the water content as a percentage.

5.2.5.2 DSC of solvent-free liquid [C-Mb][S₇]

TGA of solvent-free liquid [C-Mb][S₇] displayed a decomposition temperature of 400 °C and a water content of 0.43% (Figure 5.10). The thermal decomposition temperature of lyophilized Mb is 345 °C, displaying an increased thermal stability of the solvent-free liquid [C-Mb][S₇] construct of 55 °C, again being comparable to the BSA and PK analogues. The water content of 0.39 % for solvent-free liquid [C-Mb][S₇] was calculated; the total mass of a solvent-free liquid [C-Mb][S₇] construct was calculated to be 65970 g.mol⁻¹; the mass of cationized Mb was 20106 g.mol⁻¹ and the mass of 45864 S₇ surfactant molecules being 45864 g.mol⁻¹, yielding 14 water molecules that are highly associated with the protein construct.

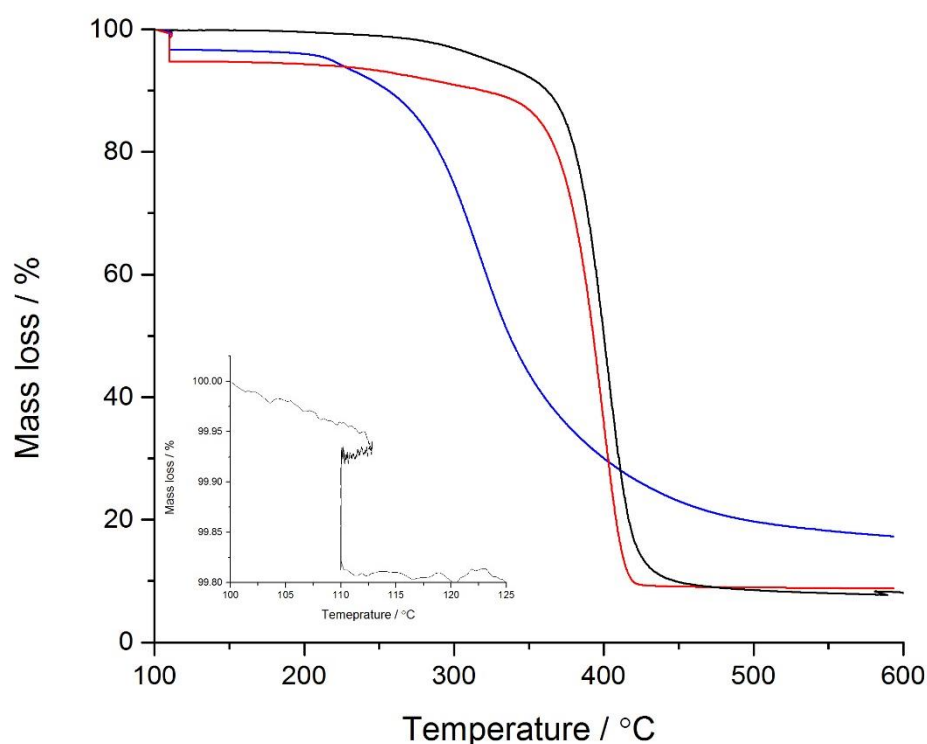


Figure 5.10: Thermogravimetric analysis of solvent-free liquid [C-Mb][S₇] (black), lyophilized Mb (blue), and S₇ surfactant (red), showing percentage mass loss in relation to temperature. The inset displays a magnified region of solvent-free liquid [C-Mb][S₇] in the range 100 - 120 °C showing the mass loss associated with the loss of surface-bound water (temperature held at 110 °C for 1 hour), *i.e.* the water content as a percentage.

5.2.6 Secondary structure analysis of solvent-free liquid myoglobin and its aqueous precursors

It is evident from the 25 ° CD spectra of native, cationized, and surfactant conjugated myoglobin, along with the solvent-free liquid myoglobin, that each surface modification altered the secondary structure (Figure 5.11). This is noted in previous literature¹⁷, however, the specific effect of each stage did differ slightly from how it was previously noted. The cause for this disparity could have been for several reasons, *e.g.* different batches of myoglobin were used and neither the previous literature nor the work conducted here displayed 100% cationization meaning it is possible for different sites to be cationized. However, the general trends described in the literature did remain despite the numerical differences in secondary structure percentage alteration.

Cationization of myoglobin to form C-Myo decreased the α -helical and β -sheet content and increased the random coil content (Table 5.1). This is displayed in the CD spectrum by a large decrease in intensity of the 195 nm characteristic wavelength associated with the $\pi \rightarrow \pi^*$ transition perpendicular to the axis of the helix and also the $\pi \rightarrow \pi^*$ transition associated with β -sheets, and a shallowing of the 222nm characteristic peak associated with the $n \rightarrow \pi^*$ transition of both the α -helices and the β -sheets (Figure 5.11).

Surfactant conjugation of C-Myo to form aqueous [C-Myo][S₇] increased the α -helical content to one comparable with the native, a larger increase than observed previously, however, the β -sheet content did decrease further, a trend not observed previously (Table 5.1). The main evidence for this in the CD spectrum was a large blue-shift (210-206 nm) and increase in the intensity of the characteristic peak relating to the negative $\pi \rightarrow \pi^*$ transition that is parallel to the axis of the helix (Figure 5.11). This showed that, as shown previously, surfactant conjugation has positive but also negative effects on the secondary structure of the protein. Addition of surfactant had the effect of screening the repulsive like charges but did not repair the salt bridges and disulfide bonds that were disrupted during cationization.

The formation of solvent-free liquid decreased the α -helical, reclaimed a percentage of β -sheet content, and retained the same turn content relative to aqueous [C-Myo][S₇]. This returned the secondary structure percentages to similar values to that of C-Mb, with the α -helical content being 45% that of the native, similar to previous work¹⁷. The main shift in the CD spectrum was a general decrease in intensity of each characteristic peak. The cationized and solvent-free variants of myoglobin from previous work were also shown to be comparable in secondary structure content percentages, reasoned through the formation of a similar molten globular-like state from both surface modifications.

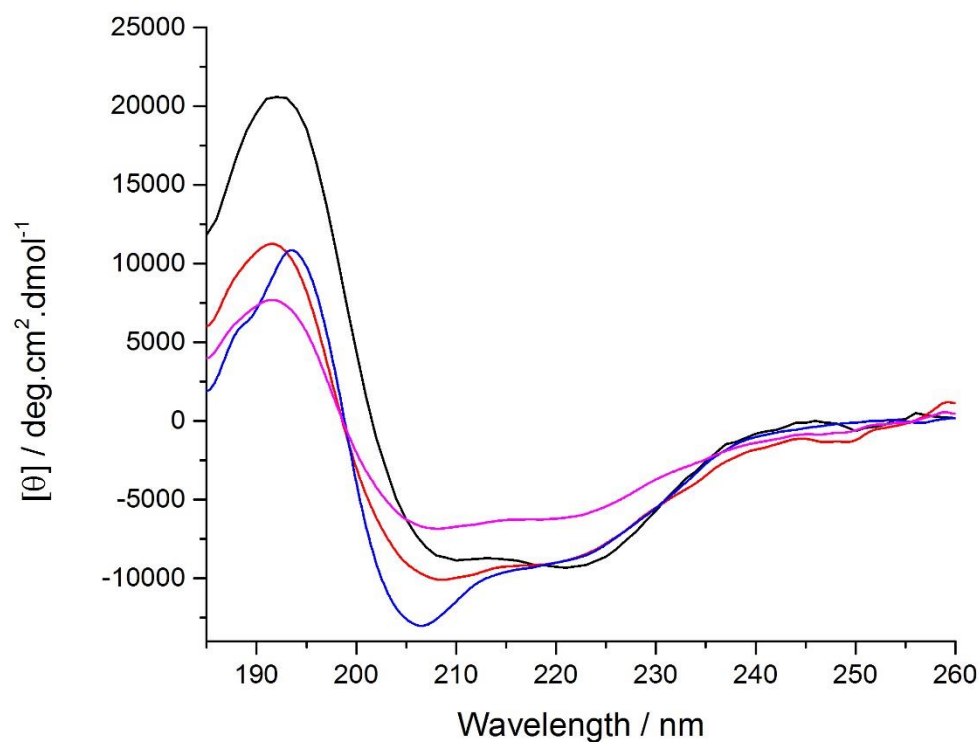


Figure 5.11: Mean residue ellipticity against wavelength showing far-UV CD spectra for of aqueous Myo (black), C-Myo (red), and aqueous [C-Myo][S₇] (blue) at 25 °C, and solvent-free liquid [C-Myo][S₇] (magenta) at 35 °C, all between 260-190 nm.

Table 5.1: Table of secondary structure content calculated from CD spectra displayed in figure.... Calculation detailed in Section 2.1.3.

Sample	α -Helix / %	β -Sheet / %	Turns / %	Other/ %
Solvent-free liquid [C-Myo][S ₇]	10	21	14	55
Aqueous [C-Myo][S ₇]	24	11	15	50
Aqueous C-Myo	11	23	14	52
Aqueous native-Myo	22	30	12	36

5.2.7 UV/Vis spectroscopy of myoglobin

UV/Vis spectroscopy is a common technique used to study proteins that contain a prosthetic group such as myoglobin (myoglobin contains a prosthetic protoporphyrin IX group that has an iron centre (haem)). The haem group is connected to the myoglobin within haem pocket as a ligand to the iron centre *via* a proximal histidine (His-93). The haem within myoglobin has a specific set of characteristic absorbances in the UV spectrum; a sharp Soret band ($\lambda = 409$ nm) and two associated Q bands ($Q_{\alpha} = 504$ nm and $Q_{\beta} = 535$ nm) all related to $\pi \rightarrow \pi^*$ transitions²⁰. The wavelength of these bands is dependent on the conformation state of myoglobin, the ligand that occupies the sixth position on the iron centre, and on the oxidation of the iron centre. As such, any deviation from the native environment will be observed as a red or blue shift in the peaks of these characteristic bands, with the direction and magnitude of the shift being used to investigate the modified environment. Previous investigations on solvent-free liquid myoglobin and its aqueous precursors have shown that the surface modifications performed on myoglobin do alter the characteristic peak positions of both the Soret and Q bands (Table 5.2)¹⁷. These deviations from a native-like absorbance spectrum are due to alterations in the dielectric constant and conformation originating from the surface modifications along with the removal of water potentially altering the haem binding geometry²¹. As such, these experiments were replicated to ensure that the solvent-free liquid myoglobin synthesised is analogous to previous literature work (Table 5.3)¹⁷.

It is evident from the alterations in spectra and specific shifts in the characteristic haem peaks that the surface modifications performed altered the local haem environment (Figure 5.12) and yielded similar peak characteristic wavelengths as previous work (Table 5.2 and Table 5.3). Whilst there were small deviations in the myoglobin aqueous precursor spectra, relative to previous work, the solvent-free [C-Myo][S₁] samples yielded identical peak positions.

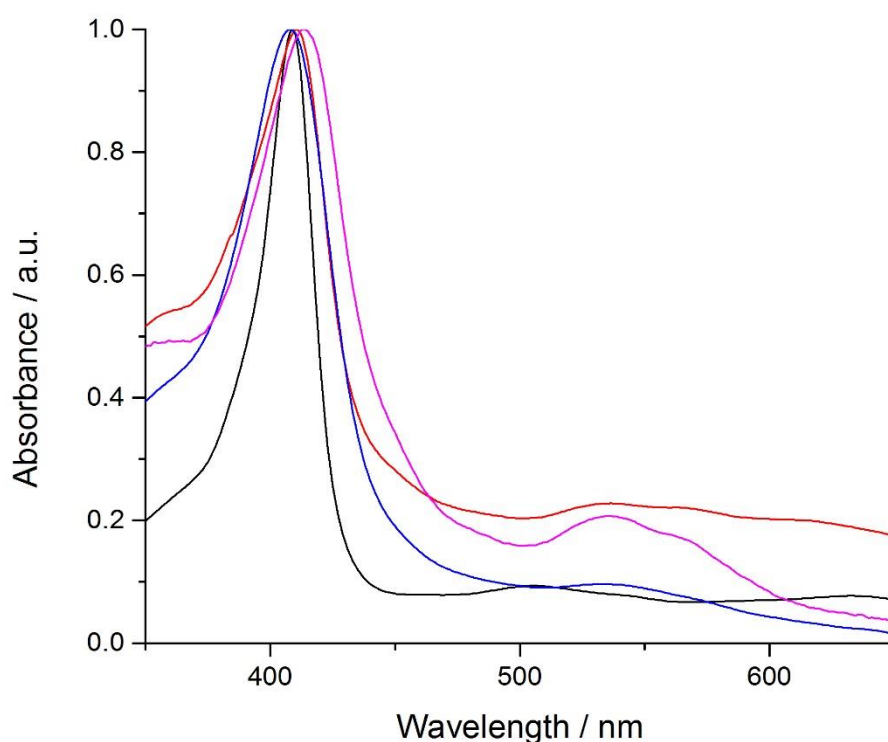


Figure 5.12: UV/Vis spectra showing the prosthetic haem group Soret band and Q absorptions for aqueous (pH 6.5 – 7, 0.2 mg.mL⁻¹) myoglobin (black), C-Myo (red), [C-Myo][S₁] (blue), and solvent-free liquid [C-Myo][S₁] (magenta). All spectra were normalized to the Soret band absorption. Experiments were conducted at 25 °C and the displayed spectra are all averages of at least 3 runs.

Table 5.2: Table displaying peak UV/Vis absorbances for solvent-free [C-Myo][S₁] and its aqueous precursors from literature¹⁷.

	Soret band (nm)	Q _α band (nm)	Q _β band (nm)
Myo	409	504	540
C-Myo	410	536	N/A
Aqueous	412	535	566
Solvent-free [C-Myo][S ₁]	414	536	564

Table 5.3: Table displaying peak UV/Vis absorbances for solvent-free [C-Mb][S₁] and its aqueous precursors from this thesis.

	Soret band (nm)	Q _α band (nm)	Q _β band (nm)
Myo	409	504	536
C-Myo	410	537	564
Aqueous	408	532	N/A
Solvent-free [C-Myo][S ₁]	414	536	564

5.2.8 Conclusion of characterisation

In conclusion, through analysis by numerous different characterisation techniques, it can be stated that the solvent-free liquid [C-Myo][S_x] synthesised (where x = 1 or 7) is analogous to that in previous literature^{17,18}. As such, the use of solvent-free liquid [C-Myo][S₁] in the formation of colloidosomes containing high concentrations of modified protein can proceed.

5.3 Results: Colloidosome formation

Colloidosomes containing high concentrations of aqueous [C-BSA][S₁] nanoconjugate were created using the method detailed in Section 2.3.2.3. Size analysis was performed on the samples prior to and post drying and transferal to water to investigate any potential size changes due to the process. A fluorescent confocal microscopic study was carried out prior to and post-drying and transferal to water to ascertain the location of the stabilizing silica and the encapsulated protein, elucidating whether the protein is retained within the colloidosome or if some protein leakage occurs.

5.3.1 Size distribution

Colloidosomes were successfully formed encapsulating surfactant conjugated protein (Figure 5.13). Optical microscopy images of colloidosomes samples were examined using image analysis software. This was conducted on colloidosomes samples pre-drying and post-drying and transferal to water to investigate any alteration to the average size and the distribution of sizes within a colloidosomes population that these processes had. If the silica shell is crosslinked to a high degree then it should present as rigid, the removal of the external oil phase and internal water phase not altering the size or shape of the colloidosome. However, if it has not crosslinked to a high enough degree, then the size and shape of the colloidosome may alter during drying. If a small degree of buckling is present, to the extent where the colloidosome is still approximately spherical, then this would increase the surface area to volume ratio and give the colloidosome an additional RBC characteristic.

The calculated average diameter of colloidosomes pre-drying was $97 \pm 54 \mu\text{m}$ and post-drying and transferal to water was $88 \pm 45 \mu\text{m}$ (Figure 5.14). The indicates that general distribution of diameters does not differ to a large extent. Both sets of colloidosomes have a very polydisperse population, with polydispersity indices of 0.53 and 0.56 for the pre-drying sample and the post-drying and transferal to water sample respectively, again displaying only a small difference between the two.

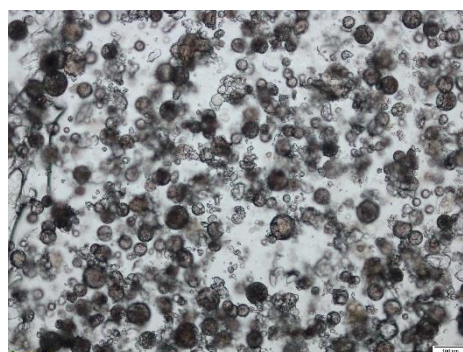


Figure 5.13: An optical microscopy image (scale bar = 100 μm) of silica colloidosomes containing aqueous [C-BSA][S₁] nanoconjugate dissolved in water (concentration 20mg/50 μl) before drying.

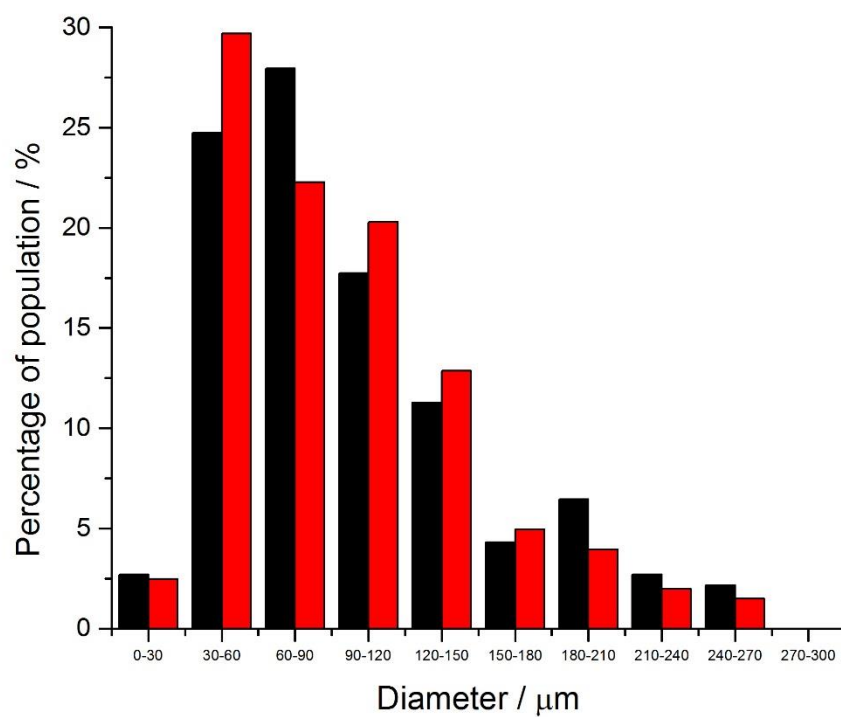


Figure 5.14: Histogram displaying the size distribution of myoglobin colloidosomes A. pre-drying (black) and B. post-drying and transfer to water (red) as percentages of the population.

5.3.2 Location of silica and encapsulated protein

5.3.2.1 Encapsulating FITC-tagged BSA nanoconjugate

Once it had been demonstrated that colloidosomes could be formed using a highly concentrated dissolved solution of solvent-free liquid [C-BSA][S₁], it was necessary to demonstrate that the protein had been encapsulated within the colloidosome and was not partially/fully dispersed within the oil phase of the emulsion, a potential outcome as discussed previously. The BSA used to create the solvent-free liquid protein was tagged with fluorescein isothiocyanate (FITC) prior to the creation of the solvent-free liquid protein (detailed in Section 2.3.2.2.1), allowing for confocal laser microscopy images to be taken.

It is clear from the confocal image (Figure 5.15A) of the pre-drying colloidosome sample that the FITC-tagged aqueous [C-BSA][S₁] nanoconjugates are encapsulated within the colloidosome and are not dispersed within the oil phase. This image also displays the inhomogeneity of the encapsulated BSA, implying that it is not only BSA that is encapsulated, as the FITC-tagged BSA would be dispersed evenly throughout. The aggregates within the colloidosome are thought to be small aggregates of silica and cause of their green fluorescence is due to the FITC-tagged BSA strongly interacting with the silica²². The origin of the internal silica is either from the TMOS silicification or from the particles of silica that were used to form the Pickering emulsion.

From the Z-stack image of the pre-drying colloidosome (Figure 5.15B) it is evident that a near-spherical colloidosome is formed. It was hypothesised that the marginally non-spherical nature arises from interactions with neighbouring colloidosomes or free silica aggregates in the sample that are present during silicification but later detach, leaving a depression in the crosslinked shell. The other potential cause for deformation is water evaporation during the timer period of crosslinking, slightly reducing the volume prior to the solidification of the shell.

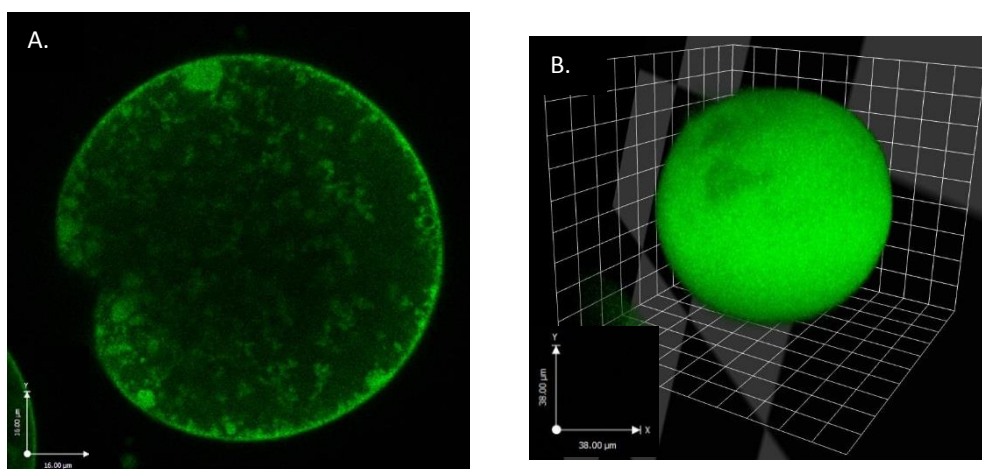


Figure 5.15: **A.** Confocal microscopy image (scale bar = 16 μm) and **B.** Z-stack image (scale bar = 38 μm), of silica colloidosomes containing aqueous FITC-tagged [C-BSA][S₁] nanoconjugate dissolved in water (concentration 20mg/50μl) before drying.

Once the colloidosome has been dried and subsequently transferred to water, the internal nature of the colloidosome changes; rather than being mostly homogenous throughout, the FITC-tagged BSA nano-conjugate is concentrated to the walls of the colloidosome and within the silica aggregates present within the colloidosome (Figure 5.16A). This indicates that during the drying stage of the transfer process, the FITC-tagged BSA interacts strongly with the silica within the wall of the colloidosome, with re-suspension within water not re-dissolving the protein. Post-transfer, the majority of the colloidosomes retain their near-spherical morphology (Figure 5.16 A and C), however, a small percentage of the colloidosomes drastically deform (Figure 5.16B). This implies that the crosslinked shell formed around a small percentage of the colloidosomes is not as rigid as the rest.

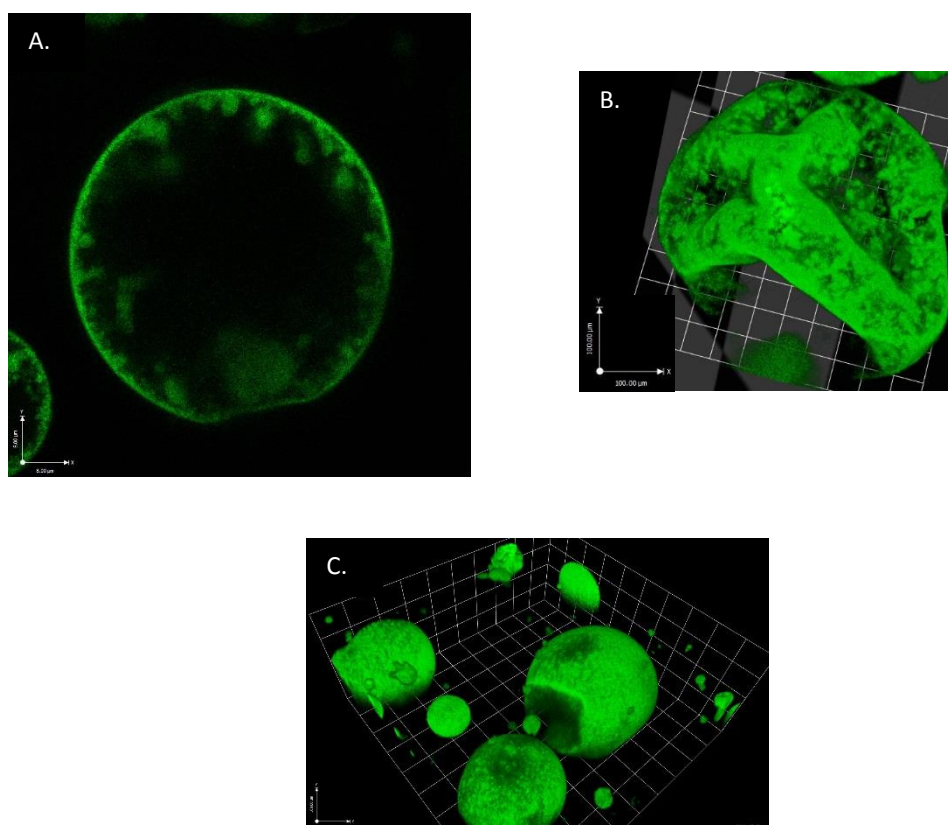


Figure 5.16: **A.** and **B.** Confocal microscopy images (scale bar = 8 μm) and **C.** a Z-stack image (scale bar = 100 μm), of silica colloidosomes containing aqueous [C-BSA][S₁] nanoconjugate dissolved in water (concentration 20mg/50 μl) after drying and transferal to water .

5.3.2.2 Encapsulating FITC-tagged BSA nanoconjugate with RITC-tagged silica

To investigate the origin of the two different populations of silica within the colloidosome sample (that which is from the original stabilization and that which is from the crosslinking TMOS) the silica used to initially stabilize the Pickering emulsion was tagged with rhodamine isothiocyanate (RITC). A successful initial control was conducted to ensure that the doped silica didn't affect the formation of colloidosomes (Figure 5.17), with near-spherical water-containing colloidosomes being formed. It is important to note that the internal fluorescence is only marginally higher than that of the background implying that there is very little of the stabilizing silica present. This indicates that the internalization

of silica could be due to the alteration of surface tensions due to the presence of the surfactant conjugated protein.

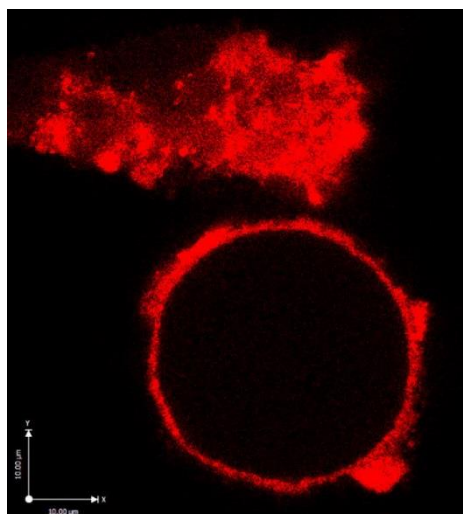


Figure 5.17: Confocal microscopy image with the RITC channel open (scale bar = 10 μm) of RITC-tagged silica colloidosome before transfer to water.

Colloidosomes encapsulating FITC-tagged BSA nano-conjugates created using doped silica for the initial Pickering emulsion stabilization were formed (Figure 5.18). It was evident that it wasn't only the FITC-tagged BSA that was within the water phase of the colloidosome; excess doped silica also pervaded the colloidosome's interior (Figure 5.18B), seen here as orange due to the merging of the coloured images. From the z-stack images it is again clear that there is excess doped silica surrounding the formed colloidosomes (Figure 5.18A).

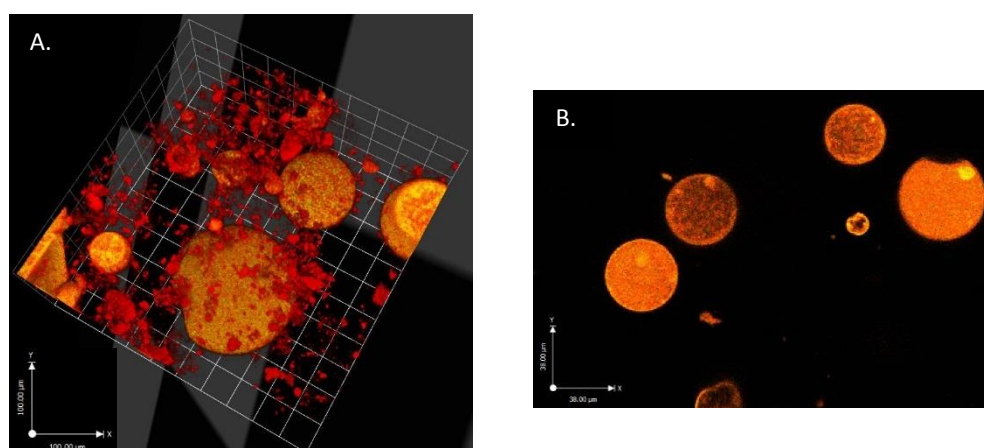


Figure 5.18: **A.** z-stack confocal microscopy image (scale bar = 100 μm) and **B.** confocal microscopy image with RITC channel open (scale bar = 100 μm) of RITC-tagged silica colloidosomes before drying containing solvent-free liquid FITC-tagged BSA dissolved in water (concentration 20mg/50 μl).

Transfer of the colloidosome to water displayed that empty or partially empty colloidosomes remained, the partially empty colloidosomes containing silica aggregate as displayed previously (Figure 5.19). As with the non-doped colloidosomes post transfer to water, the amount of excess silica present is vastly reduced due to its aggregation and rapid sedimentation onto the sample slide. Again, the transfer process is damaging to a small percentage of the colloidosomes, with shell breakages and occasional full fractures. Another observation is the high retention of BSA-nanoconjugate in the post-transferral colloidosomes, displaying again that water does not re-dissolve the nano-conjugate once it has been in contact with silica (due to different microscope settings the merging of the green and red channels appears yellow rather than orange, displaying where both the FITC-tagged BSA nanoconjugates and the doped silica are located). If the water did re-dissolve the protein, any cracks within the shell wall would therefore leak the protein. This is key to the retention of myoglobin/haemoglobin nano-conjugate and for the formation of a synthetic red blood cell.

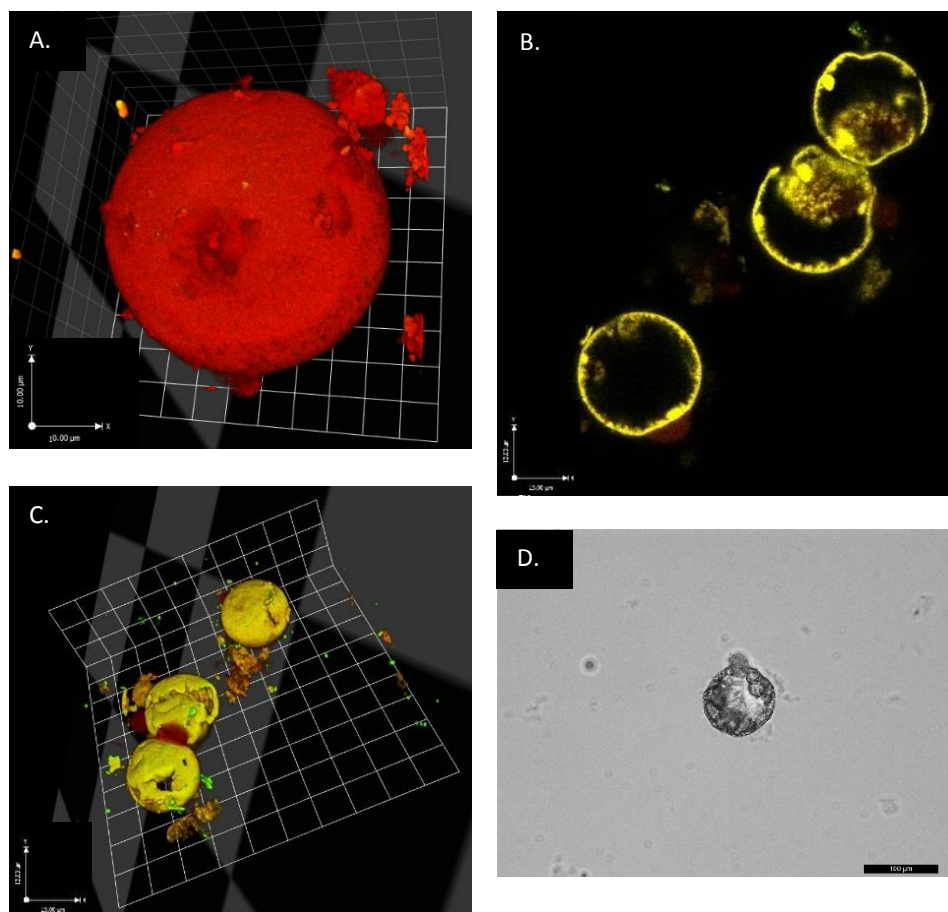


Figure 5.19: **A.** Z-stack image (scale bar = 100 μm), **B.** confocal microscopy image (scale bar = 38 μm), **C.** Z-stack image with both the FITC and RITC channel open (scale bar = 38 μm), and **D.** an optical microscopy image (scale bar = 100 μm) of RITC-tagged silica colloidosomes after drying and transferal to water that contained solvent-free liquid FITC-tagged BSA dissolved in water (concentration 20mg/50 μl before drying).

5.3.3 Encapsulation of myoglobin nano-conjugate

As has been discussed in Section 1.2.5 and displayed in Chapter 3, each solvent-free liquid protein formed has similar characteristics and therefore the solvent-free liquid FITC-tagged BSA can be replaced with solvent-free liquid myoglobin in the formation of colloidosomes. Myoglobin was selected due to its ability to bind and de-bind gaseous molecules, namely oxygen, and because myoglobin had previously been used to create a solvent-free liquid myoglobin which has been extensively studied^{17,18}. Myoglobin is used here as a preliminary substitute for haemoglobin to examine the ability of a high concentration of encapsulated protein to bind, retain, and release oxygen to mimic a red blood cell.

As is clear from confocal images of colloidosomes pre and post transfer to water (Figure 5.20 and Figure 5.21 respectively) the colloidosomes formed are of a similar near-spherical nature, internally displaying near-homogeneity pre-transfer and a similar internal and external structure post-transfer to the colloidosomes encapsulating FITC-tagged BSA nano-conjugate. It can be assumed that the myoglobin nano-conjugate is in similar locations to the BSA counterpart, strongly interacting to the internal silica aggregates and to the silica shell.

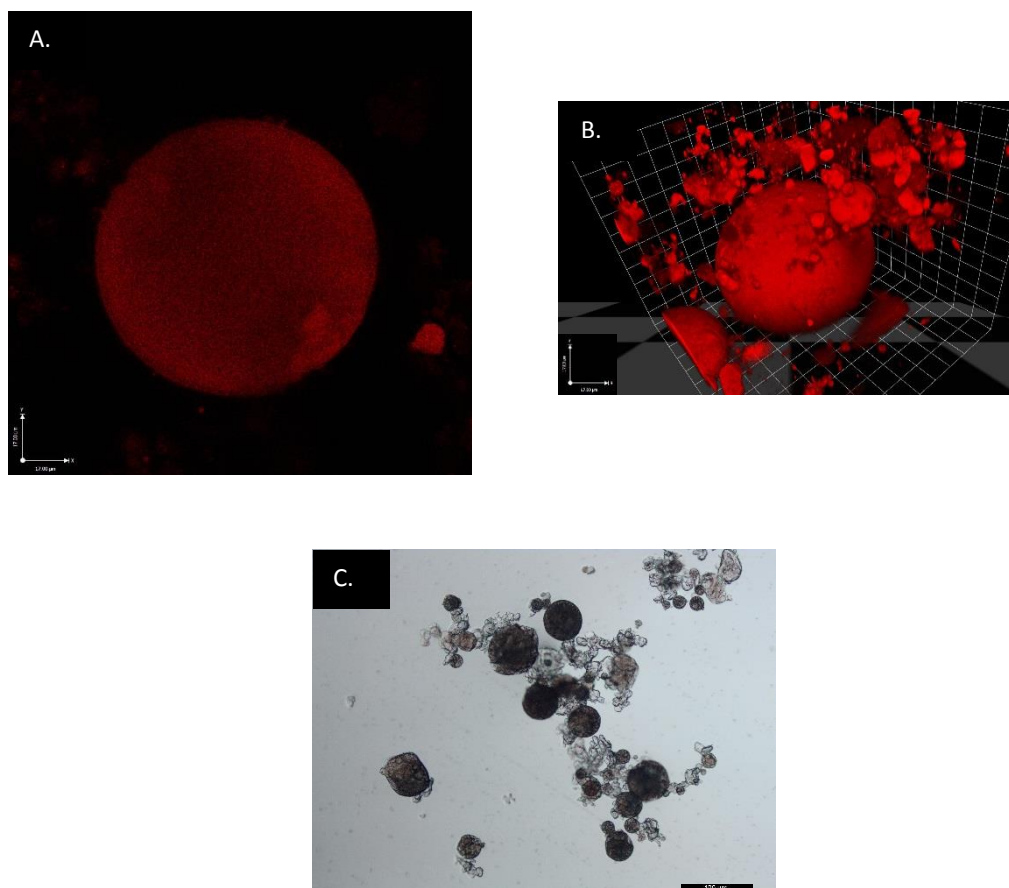


Figure 5.20: **A.** Confocal microscopy image (scale bar = 17 μm), **B.** Z-stack image (scale bar = 17 μm), and **C.** an optical microscopy image (scale bar = 100 μm) of RITC-tagged silica colloidosomes before drying containing solvent-free liquid [C-Mb][S₁] dissolved in water (concentration 20mg/50μl).

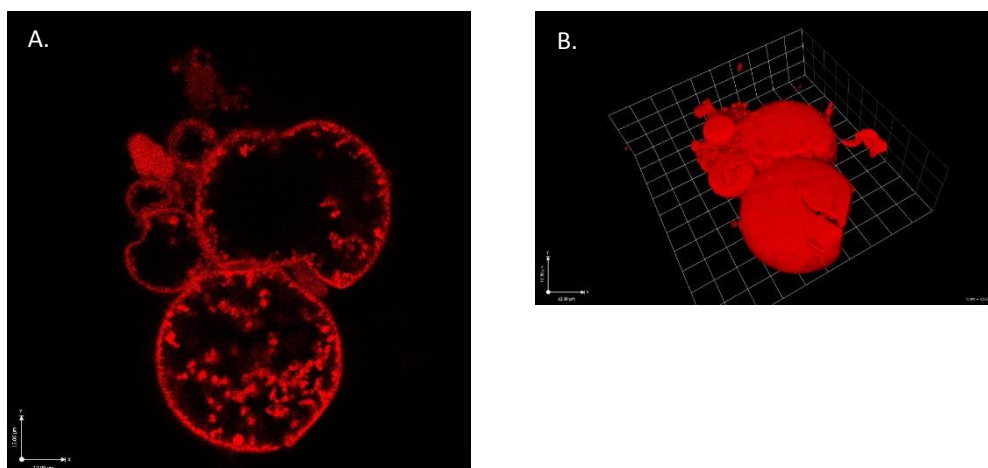


Figure 5.21: **A.** Confocal microscopy image (scale bar = 12 μm) and **B.** Z-stack image (scale bar = 12 μm), of RITC-tagged silica colloidosomes after drying and transferal to water that contained solvent-free liquid [C-Mb][S₁] dissolved in water (concentration 20mg/50 μl before drying).

5.3.3.1 Concentration of myoglobin within colloidosomes

The starting concentration of myoglobin within the aqueous phase of the Pickering emulsion was 247 mgmL^{-1} . As there was no evidence of protein loss when using FITC-tagged BSA nanoconjugate, *i.e.* no fluorescence outside of the colloidosome, it is assumed that there is also no loss of the myoglobin nanoconjugate. As such, the final concentration of the myoglobin nanoconjugate within the colloidosomes is approximately 247 mgmL^{-1} , over 12 times higher than that attainable with none modified myoglobin (20 mgmL^{-1}).

5.3.4 UV-Vis and DR UV-Vis

As discussed previously in Section 5.2.7, the haem contained within myoglobin displays characteristic peaks in a UV/Vis spectrum, predominantly a Soret band at 409 nm (when in its native form). The surface modifications required to form solvent-free liquid [C-Mb][S₁] were shown to alter the haem's local environment, shifting the characteristic peak positions. This same method can be utilised to confirm the presence of myoglobin within the colloidosomes; the presence of a Soret band only being able to originate from the myoglobin and not the silica. However, standard aqueous UV/Vis spectrometry is not applicable to the use of colloidosomes due to the light scattering effects generated by solid opaque objects, *i.e.* colloidosomes. This causes the spectrum to appear effectively flat with no characteristic peaks present as very little to no light is detected (Figure 5.22, black line) even whilst using low volume low path length cuvettes. Diffuse reflectance UV/Vis (DR-UV/Vis) was used in its place, a technique that detects reflected light from a sample. For this, the colloidosome samples were dried down to a powder through lyophilization and placed between two glass slides.

As is evident from the DR-UV/Vis spectrum, the Soret band is present (406 nm), however, both of the Q bands are too weak to be discerned (Figure 5.22, green line). The presence of the Soret band in DR-UV spectrum indicates the presence of myoglobin within formed colloidosomes

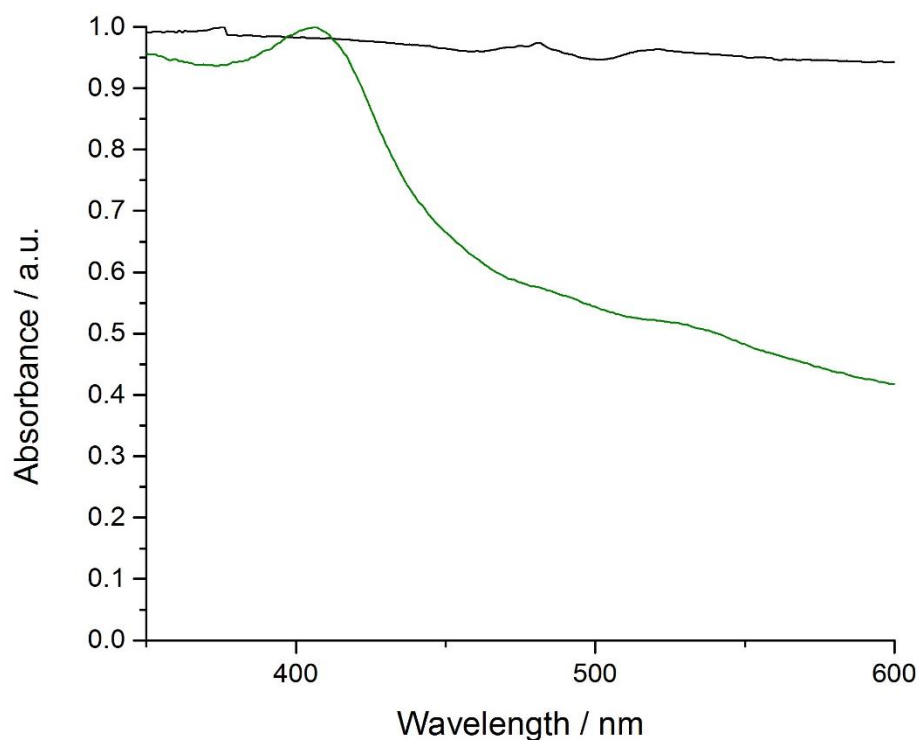


Figure 5.22: UV/Vis spectra showing the prosthetic haem group Soret band for aqueous [C-Mb][S1] nanoconjugate (black), and DR-UV/Vis spectra for dried [C-Mb][S1] nanoconjugate (green). Both spectra were normalized. Experiments were conducted at 25 °C and the displayed spectra are all averages of at least 3 runs.

5.3.5 Scanning electron microscopy (SEM) of colloidosomes

SEM was used to image the colloidosomes that encapsulated aqueous [C-Mb][S₁] nanoconjugate providing highly detailed images of the colloidosomes, displaying that the colloidosomes are rigid structures and that a general spherical geometry is retained under ultra-high vacuum.

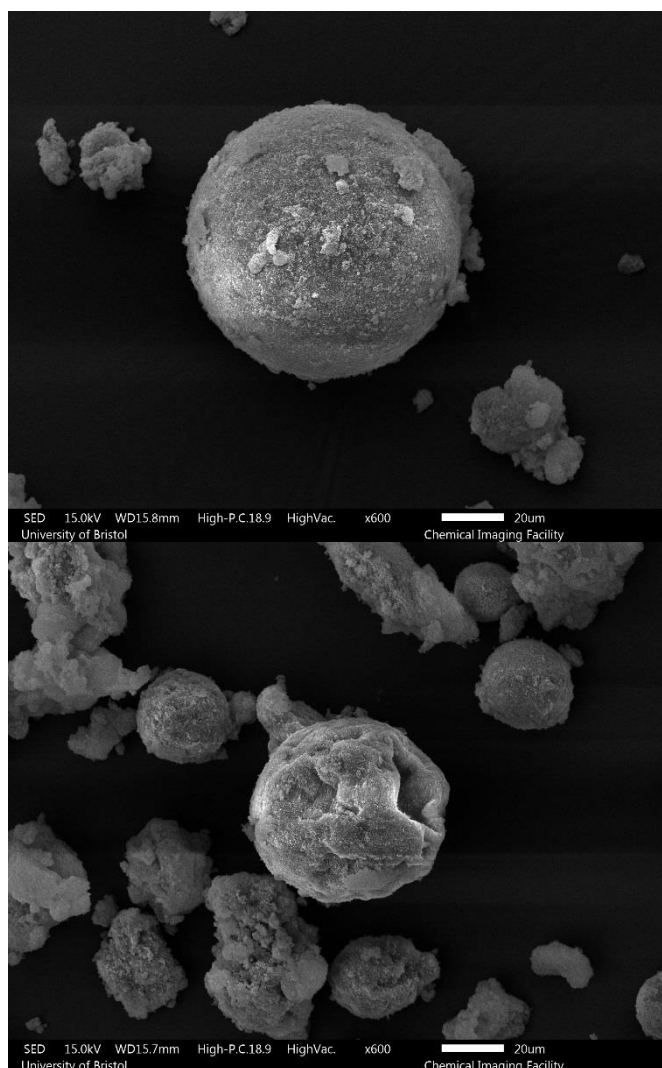


Figure 5.23: Scanning electron microscopy images of colloidosomes containing high concentrations of [C-Mb][S₁] nanoconjugate. The scale bar displays 20 μ m

5.4 Conclusions and future work

The encapsulation of a highly concentrated protein nano-conjugate solution within a silica crosslinked colloidosome permits the creation of a synthetic red blood cell (sRBC) containing an as up to now unattainable concentration of myoglobin. Prior to transferal to water, the protein was shown to be internally near-homogeneous within the colloidosome, as well as interacting with the shell. After transfer to water, the protein was shown to be solely located within the shell of the colloidosome, displaying a strong interaction with the silica to the extent of not being re-dissolved within the water-phase. The substitution of solvent-free liquid [C-BSA][S₁] with solvent-free liquid [C-Mb][S₁] was shown to not affect the morphology of the colloidosome. Once dried, DR-UV/Vis was used to definitively show the presence of myoglobin within the colloidosomes. However, from DR UV-Vis analysis of the dry colloidosomes, the encapsulated myoglobin has retained a native-like structure and it is therefore predicted that the myoglobin will possess the desired function.

Due to instrumental limitations the analysis of encapsulated myoglobin oxygen binding, retention, and release was unable to be conducted. To display that the encapsulated [C-Mb][S₁] nanoconjugate is functional as an oxygen binder/de-binder, an experimental set-up as detailed in the literature is required, with a DR-UV/Vis spectrometer being within a highly controlled air environment (a glove box)¹⁸. The myoglobin would be thoroughly de-oxygenated, shifting the Soret and Q bands due to an alteration to the haem's local environment. If the myoglobin retains its functionality, when oxygen is added to the system these bands will again shift as the local environment alters.

Once myoglobin has been shown to be functionally present once encapsulated, it can be substituted with haemoglobin, thus forming a novel version of a synthetic red blood cell that displays high concentrations of functional haemoglobin. The high concentration could then be combined with other sRBC techniques^{14,15}, creating a sRBC that more realistically represents a natural RBC.

5.5 References

1. Nedovic, V., Kalusevic, A., Manojlovic, V., Levic, S. & Bugarski, B. An overview of encapsulation technologies for food applications. *Procedia Food Sci.* **1**, 1806–1815 (2011).
2. Binks, B. P. & Clint, J. H. Solid wettability from surface energy components: Relevance to pickering emulsions. *Langmuir* **18**, 1270–1273 (2002).
3. McGorty, R., Fung, J., Kaz, D. & Manoharan, V. N. Colloidal self-assembly at an interface. *Mater. Today* **13**, 34–42 (2010).
4. Chevalier, Y. & Bolzinger, M. A. Emulsions stabilized with solid nanoparticles: Pickering emulsions. *Colloids Surfaces A Physicochem. Eng. Asp.* **439**, 23–34 (2013).
5. Bon, S. A. F. CHAPTER 1. The Phenomenon of Pickering Stabilization: A Basic Introduction. 1–7 doi:10.1039/9781782620143-00001
6. Dinsmore, A. D. *et al.* Colloidosomes: Selectively permeable capsules composed of colloidal particles. *Science (80-.)*. **298**, 1006–1009 (2002).
7. Brinker, C. J. Hydrolysis and Condensation of Silicates: Effects on Structure. *J. Non. Cryst. Solids* **100**, 31–50 (1988).
8. Buckley, A. & Greenblatt, M. The sol-gel preparation of silica gels. *J. Chem. Educ.* **71**, 599–602 (1994).
9. Li, M., Huang, X. & Mann, S. Spontaneous growth and division in self-reproducing inorganic colloidosomes. *Small* **10**, 3291–3298 (2014).
10. Huo, C., Li, M., Huang, X., Yang, H. & Mann, S. Membrane engineering of colloidosome microcompartments using partially hydrophobic mesoporous silica nanoparticles. *Langmuir* **30**, 15047–15052 (2014).
11. Li, M., Harbron, R. L., Weaver, J. V. M., Binks, B. P. & Mann, S. Electrostatically gated membrane permeability in inorganic protocells. *Nat. Chem.* **5**, 529–536 (2013).
12. Li, M., Green, D. C., Anderson, J. L. R., Binks, B. P. & Mann, S. In vitro gene expression and enzyme catalysis in bio-inorganic protocells. *Chem. Sci.* **2**, 1739 (2011).
13. Patwardhan, S. V & Clarson, S. J. Silicification and biosilicification - Part 7: Poly-L-arginine mediated bioinspired synthesis of silica. *J. Inorg. Organomet. Polym.* **13**, 193–203 (2011).
14. Diez-Silva, M., Dao, M., Han, J., Lim, C.-T. & Suresh, S. Shape and Biomechanical Characteristics of Human Red Blood Cells in Health and Disease. *MRS Bull.* **35**, 382–388 (2010).
15. Pedro, C. & Marcos, I. BLOOD SUBSTITUTES: EVOLUTION FROM NON-CARRYING TO OXYGEN AND GAS CARRYING FLUIDS. *ASAIO J* **59**, 337–354 (2013).
16. HF, B. *Approach to the anemias*. (Elsevier Saunders, 2015).
17. Brogan, A. P. S., Siligardi, G., Hussain, R., Perriman, A. W. & Mann, S. Hyper-thermal stability and unprecedented re-folding of solvent-free liquid myoglobin. *Chem. Sci.* **3**, 1839 (2012).
18. Perriman, A. W. *et al.* Reversible dioxygen binding in solvent-free liquid myoglobin. *Nat. Chem.* **2**, 622–6 (2010).
19. Hersleth, H.-P., Hsiao, Y.-W., Ryde, U., Görbitz, C. H. & Andersson, K. K. The crystal structure of peroxymyoglobin generated through cryoradiolytic reduction of myoglobin compound III during data collection. *Biochem. J.* **412**, 257–264 (2008).
20. Wiwatchaiwong, S., Nakamura, N. & Ohno, H. Spectroscopic Characterization and Electrochemistry of Poly (ethylene oxide) -Modified Myoglobin in Organic Solvents. 1276–

1281 (2006).

21. Perriman, A. W. *et al.* Reversible dioxygen binding in solvent-free liquid myoglobin. *Nat. Chem.* **2**, 622–6 (2010).
22. Kumar, S., Aswal, V. K. & Kohlbrecher, J. SANS study of Lysozyme vs. BSA protein adsorption on silica nanoparticles. *AIP Conf. Proc.* **1447**, 181–182 (2012).

Chapter 6 Conclusions

6.1 Conclusions

Novel solvent-free liquid proteins were synthesised of both proteinase K (PK) and *bovine serum albumin* (BSA) through protein surface re-engineering. Thermogravimetric analysis (TGA) of the liquids demonstrated that fewer than 30 water molecules were present, 2-3 orders of magnitude lower than the number required for aqueous protein solvation. Differential scanning calorimetry (DSC) determined that the liquids had distinct melting and crystallization temperatures which differed from those of pure surfactant, indicating that protein-surfactant nano-constructs were created. Additionally, circular dichroism spectroscopy (CD) showed that the conformations of the proteins in the solvent-free liquid state were similar to the respective native conformations, with both proteins retaining high levels of native-like secondary structure. Furthermore, through the development of a novel and innovative CD-based enzyme activity assay, protease activity was shown to be retained in the absence of solvent, exhibiting that the dynamic freedom required for proteolytic activity was retained.

The initial stage of the solvent-free liquid protein synthesis was chemical cationization through the amine displaying molecules to the surface of the protein *via* carbodiimide activation, creating cationized PK (C-PK) and cationized BSA (C-BSA). This was required to increase the number of positive point charges on the surface and reduce negative charges. Increasing the number of positive charges increases the maximum number of anionically charged surfactant molecules that can be electrostatically grafted to the surface, subsequently increasing the likelihood for the formation of a solvent-free liquid protein. Removing negative charge from the surface also reduces the like-charge repulsion. Mass spectrometry was used to calculate the number of cationized sites, with Protein Data Bank data of both proteins used to calculate the number of available sites. Cationization efficiencies for BSA and PK were calculated to be 100% and 67% respectively, therefore, permitting a maximum of 181 and 38 anionic surfactant molecules to be grafted to the surface of the proteins, respectively.

Dynamic light scattering (DLS) was used to analyse the size distribution of the proteins throughout the synthesis. Cationization of BSA was shown to increase the average hydrodynamic radius by over 20 nm, an indication of protein aggregation. C-PK did not form aggregates, showing that the aggregation of C-BSA was not due to the protein having a more positive charge. Instead, the aggregation was attributed to the partial-unfolding of BSA once cationized, exposing the hydrophobic core residues to the solvent leading to inter-protein β -sheet formation. Surfactant conjugation onto C-BSA, forming surfactant conjugated C-BSA (aqueous [C-BSA][S]), disassembled the C-BSA aggregates, with the final radii of the aqueous [C-BSA][S] being less than 1 nm larger than native-BSA. This was hypothesised to occur due to charge screening of the additional repulsive surface charges, internalising the exposed hydrophobic residues. The formation of aqueous [C-PK][S] nano-conjugates also increased the radius by a similar distance relative to native-PK.

Table 6.1: Table summarising the secondary structure and calculated thermodynamic values for solvent-free liquid [C-BSA][S₇] and Solvent-free liquid [C-PK][S₇] as well as all of the aqueous precursors for both proteins.

Sample	α -Helix / %	β -Sheet / %	Turns / %	Other / %	T _m a/ °C	ΔH_m a/ kJ mol ⁻¹	ΔS_m a/ J K ⁻¹ mol ⁻¹	T _m b/ °C	ΔH_m a/ kJ mol ⁻¹	ΔS_m a/ J K ⁻¹ mol ⁻¹
Solvent-free liquid [C-BSA][S ₇]	29	18	13	40	N/A	N/A	N/A	N/A	N/A	N/A
Aqueous [C-BSA][S ₇]	18	22	14	46	59 ± 12	102 ± 16	309 ± 46	N/A	N/A	N/A
Aqueous C-BSA	18	26	14	42	89 ± 1 *	N/A	N/A	N/A	N/A	N/A
Aqueous native-BSA	37	14	12	37	72 ± 3	320 ± 14	930 ± 9	N/A	N/A	N/A

Sample	α -Helix / %	β -Sheet / %	Turns / %	Other / %	T _m a/ °C	ΔH_m a/ kJ mol ⁻¹	ΔS_m a/ J K ⁻¹ mol ⁻¹	T _m b/ °C	ΔH_m a/ kJ mol ⁻¹	ΔS_m a/ J K ⁻¹ mol ⁻¹
Solvent-free liquid [C-PK][S ₇]	9	31	15	43	N/A	N/A	N/A	N/A	N/A	N/A
Aqueous [C-PK][S ₇]	10	26	17	46	39 ± 14	350 ± 89	1100 ± 280	59 ± 3	320 ± 12	950 ± 38
Aqueous C-PK	18	33	13	37	39 ± 4	220 ± 15	700 ± 49	61 ± 10	590 ± 73	1800 ± 220
Aqueous native-PK	10	34	13	43	33 ± 16	450 ± 160	1500 ± 500	61 ± 6	410 ± 29	1200 ± 87

CD spectroscopy played a crucial role in analysing the effect solvent-free liquid protein synthesis had on both proteins, with the spectra from each stage of the synthesis for both proteins being deconvoluted to investigate the secondary structure percentages. A CD-thermal denaturation study was conducted examining the effect on each protein's stability, a summary of the results is in Table 6.1.

The adjustment in global and local charge due to cationization of BSA initiated a large destabilization of the secondary structure, halving the α -helical content and almost doubling the percentage of β -sheets. The loss of α -helical structure can lead to the exposure of previously internalised hydrophobic core residues and lower the resistance of the protein to aggregation. The exposed residues have a propensity to aggregate, leading to the formation of inter-protein β -sheets as observed through DLS. The thermal denaturation study showed that C-BSA had become molten globular-like, not denaturing in the two-step mechanism observed for native-BSA, therefore requiring chemical denaturation. It is postulated that the formation of aggregates also influenced the thermal denaturation dynamics of the protein. Interestingly, the indirect method used to calculate the denaturation temperature of C-BSA yielded a half denaturation result higher than that of native-BSA, indicating that this loosely folded structure was more thermally stable. The cause of this higher thermal stability relative to native BSA is due to C-BSA starting off in a lower enthalpic state, therefore taking a higher energy to denature to a similar degree. Surfactant conjugation of C-BSA did not appear to alter the secondary structure of the protein to a large degree, with only a small decrease in β -sheet content, marginally returning the protein to a more native-like structure, also dissolving a large majority of the previously aggregated protein. This was a large enough change in secondary structure to also reclaim some semblance of the two state unfolding mechanism, showing that aqueous [C-BSA][S₇] have a lower thermal stability relative to native-BSA.

Cationization of PK increased the α -helical content relative to native-PK, the opposite of what occurred with C-BSA, possibly indicating that the formation of C-BSA broke several surface salt bridges and ionic pairings, whilst the formation of C-PK did not. This emphasises how different proteins change with the same surface modification, with the disparity in results being due to the variation in protein surface-chemistry and possibly the level of cationization. The thermal stability of C-PK was similar relative to native-PK, having marginally altered entropies and enthalpies of transition alongside a different structure of the stabilized thermal intermediate. Surfactant conjugation of C-PK returned the α -helical content to native-like levels but also reduced the β -sheet percentage, again displaying the prevalent structure reclamation effect of surfactant conjugation to a previously highly cationized protein. This reclamation is attributed to the surface charge shielding effect of the anionic head group, removing the additional surface repulsions originating from the cationization stage of the synthesis. The thermal stability of aqueous [C-PK][S₇] was similar relative to native-PK, although the entropies and enthalpies of the two denaturation events were both slightly lower, indicating that the modified protein has a lower thermal stability.

The formation of the two solvent-free liquids *via* lyophilization of aqueous protein-nanoconjugates and subsequent heating had varying effects on the two proteins. Solvent-free liquid [C-PK][S₇] did not show a large alteration to the secondary structure relative to aqueous [C-PK][S₇] nanoconjugates beyond a slight increase in β -sheet percentage. This signified that, in general, the solvent-free liquid protein synthesis did not have a large effect on the secondary structure content of PK.

The formation of solvent-free liquid [C-BSA][S₇] showed a large degree of structure reclamation relative to aqueous [C-BSA][S₇], bringing the α -helical and β -sheet content back to more native-like levels; an outcome attributed to the confining effect that the increased concentration has on the protein. Being in the solvent-free liquid protein state greatly increases the entropic barrier to protein unfolding, effectively compressing the protein. Therefore, the loosely bound aqueous nanoconjugate state becomes volume-restrained, making it more energetically favourable to re-form secondary structure motifs.

The thermal denaturation unfolding pathway of solvent-free liquid [C-BSA][S₇] became linear rather than sigmoidal, and as such the entropy and enthalpy of denaturation could not be calculated. However, there was a large percentage of secondary structure retained above 100 °C, signifying a considerable increase in thermal stability relative to native-BSA. The thermal denaturation unfolding pathway of solvent-free liquid [C-PK][S₇] differed from the analogous BSA construct, showing formation of a native-like intermediate species that maintained a high level of secondary structure past 100 °C. This again signifies that the solvent-free liquid version of a protein has an increased thermal stability relative to the native.

A novel CD-based enzyme assay was used to study how the surface modifications required to form a solvent-free liquid protein altered the proteolytic activity of the PK towards BSA, with the final PK/BSA pairing being the solvent-free liquid forms of each protein. The Michaelis-Menten model for enzyme kinetics was used in conjunction with the CD-assay, with values for the enzymatic rate and binding affinity being calculated (Table 6.2). Cationization reduced the enzymatic activity in all pairings, an effect attributed to the structural alteration of the protease once cationized and the increase in positive surface charge repulsion. The alteration to the structure of the protease potentially modified the conformation of the active site, changing the interaction mechanism between C-PK and the BSA substrates. An increase in the positive surface charge would have had two main effects; firstly, it increased the global repulsion between two proteins, and secondly it altered how the binding cleft and catalytic triad interacted with the substrate, with more positive charge on the substrate surface potentially reducing the strength of the interactions. Surfactant conjugation also reduced the enzymatic activity of all pairings despite the reclamation of a native-like protease structure; this effect being attributed to an increase in steric hindrance brought about from the long hydrophobic/hydrophilic tail sections of the surfactants. These tails could have hindered the binding

interactions along with any catalytic interactions between the substrate and the protease, preventing the proteolytic cleavage of the substrate.

The surface modifications also influenced how strongly the substrate and the products bound to the active site of the protease (binding affinity). Cationization of BSA decreased the binding affinity as the large surface charge increased the like-charge repulsion and increased the energy required to bind. Interestingly, cationization of PK increased the binding affinity; an outcome that is attributed to the conformational change of the active site due to cationization. In all cases, surfactant conjugation increased the binding affinity, hypothesised to be caused by the alkyl tail domains of the surfactant favourably interacting with the more unfolded products, retarding the rate of product liberation from the active site and increasing the apparent binding affinity.

Table 6.2: The maximum rate of activity (V_{\max}) and binding affinity (K_M) for all pairings of PK, C-PK, aqueous [C-PK][S], BSA, C-BSA, and aqueous [C-BSA][S].

V_{\max} ($\Delta\text{mdeg.s}^{-1}$)	NPK	C-PK	Aqueous [C-PK][S ₇]
K_M (mM)			
NBSA	128±26	35.4±2.72	1.58±0.98
	0.059±0.015	0.004±0.001	0.011±0.005
C-BSA	104±23	4.48±2.36	0.94±0.13
	0.49±0.12	0.051±0.039	0.003±0.001
Aqueous [C-BSA][S ₇]	5.86±2.42	NA	0.67±0.13
	0.024±0.014	NA	0.003±0.001

Protease activity was seen in the absence of solvent for the first time, with solvent-free liquid [C-BSA][S₇] being broken down by solvent-free liquid [C-PK][S₇]. The activity was exhibited as a large reduction to the secondary structure relative to controls, ensuring that it was not merely the protein unfolding due to being held at a higher temperature. This indicated that there was retention of the dynamical fluctuations required for proteolytic activity despite the lack of solvating water in the material. This provided further evidence towards the replication of protein-water interactions by protein-surfactant interactions. The rate of activity and the binding affinity in the solvent-free state cannot be numerically compared to the aqueous precursor PK/BSA pairings due to the experimental difficulty. However, it is likely that the activity of the protease in this solvent-free state would be lower relative to the aqueously solvated precursors due to the higher viscosity observed in previous solvent-free liquid proteins, along with a drastically decreased water availability, both decreasing the speed for the enzymatic turnover and the total number of turnovers. At the temperature utilised (50 °C) it is likely that diffusion of substrates into, and products out of, the active site through the viscous liquid would also be a limiting factor. This implied that if the temperature of analysis was raised, the decrease in viscosity could increase the rate of activity; the number of enzymatic turnovers remaining limited. As PK displayed high temperature stability when in the solvent-free liquid state, the use of

solvent-free liquid [C-PK][S₇] in higher temperature environments could now be possible. The novel CD-based protease assay developed throughout this work could also be scrutinized and then further expanded upon to bring it into common scientific usage.

The encapsulation of high concentrations of aqueous myoglobin nanoconjugate (approximately 240 mg/mL⁻¹) within silica-based colloidosomes was successful. This was preliminary work that could lead on to the formation of a synthetic red blood cell (sRBC) containing high concentrations of aqueous haemoglobin nanoconjugate. Obtaining a high concentration of haemoglobin within a sRBC is a challenge not yet completed; with other challenges such as membrane flexibility and the stabilization of synthetic haemoglobin already progressing within research. The colloidosomes, whilst being polydisperse, were stable and it was shown that upon transfer to water, the surface-modified myoglobin was located within the walls of the colloidosomes and did not enter the surrounding solvent. This research could be taken a step further by testing the oxygen binding capabilities of the encapsulated myoglobin, providing a step towards encapsulation of surface-modified haemoglobin. If the high concentration of encapsulated protein were to be successfully combined with other characteristics of a RBC, a superior model would be created and could be utilised within medicine.